

gb_smlt2.BP313835	-	82.00	138.48	158.19	91.9	BP313835	6028482017871	NCL CGA	1101
gb_smlt1.ALS597225	-	81.50	144.17	72.50	41.9	ALS597254	6020202017871	NCL CGA	1101
gb_smlt2.BG5272650	-	81.50	137.76	162.85	86.9	BG527265	606237197581	NTH_MGC	851
gb_smlt2.BP135074	-	81.00	147.37	117.96	55.1	BP135074	7094642. x1	NCL CGAP	551
gb_smlt2.BG170904	-	81.00	137.52	166.85	75.6	BG170904	602411310581	NTH_MGC	851
gb_smlt2.BG631932	-	81.00	136.41	196.12	90.8	BG631932	601276550381	NTH_MGC	851
gb_smlt1.AV664543	-	80.50	140.49	122.23	54.2	AV664543	AV664543	Bos taurus	542
gb_smlt1.AV667180	-	80.50	140.06	122.47	54.4	AV667180	AV667180	Bos taurus	544
gb_smlt2.BG3919150	-	80.50	128.52	328.04	139.0	BG3919150	6024110581	NTH_MGC	851
gb_smlt2.BG649485	-	80.50	139.40	133.55	58.6	BG649485	OV1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BG5271028	-	80.50	137.95	167.28	77.4	BG5271028	UT1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BG5271272	-	80.50	135.83	211.14	90.4	BP514272	60140288781	NTH_MGC	851
gb_smlt1.BE5741838	-	80.50	135.55	216.84	94.6	BE5741838	601047787581	NTH_MGC	851
gb_smlt2.BG3919150	-	80.50	128.52	328.04	139.0	BG3919150	6024110581	NTH_MGC	851
gb_smlt2.BG649485	-	80.00	140.26	119.68	48.1	AV664543	3165689	BANC S590	BO
gb_smlt2.AV651495	-	80.00	135.45	132.77	52.7	AV651495	HE-214-BL-B04-77A	NTH_MGC	570
gb_smlt2.BG649485	-	80.00	137.95	167.28	77.4	BG649485	UT1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BP040860	-	80.00	138.11	157.70	61.8	BP040860	OV1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BP040860	-	80.00	138.58	191.80	72.8	BP040860	OV1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BP5263738	-	80.00	136.53	193.10	73.2	BP5263738	602070579581	NTH_MGC	851
gb_smlt2.BP3376782	-	80.00	134.78	241.71	89.2	BP3376782	602070579581	NTH_MGC	851
gb_smlt2.BG649485	-	79.50	133.40	341.64	109.3	BP649485	602070579581	NTH_MGC	851
gb_smlt2.BG3919150	-	79.50	140.26	119.68	48.1	AV664543	3165689	BANC S590	BO
gb_smlt2.BG3919150	-	79.50	140.26	119.68	48.1	AV664543	3165689	BANC S590	BO
gb_smlt1.ALS314438	-	79.00	138.13	157.28	50.0	ALS314438	NC1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt1.ALS314438	-	79.00	138.09	157.99	50.2	ALS314438	NC1-H0003-x1-0031-01-5	NTH_MGC	570
gb_smlt2.BP6265655	-	79.00	133.23	294.71	86.8	BP6265655	602126126781	NTH_MGC	851
gb_smlt2.BG411938	-	79.00	133.79	311.77	91.2	BG411938	HYSMEM020C015	Bord	912
gb_smlt2.BP3050200	-	79.00	132.67	316.94	93.5	BP3050200	60188841261	NTH_MGC	851
gb_smlt2.BP3050200	-	79.00	132.67	316.94	93.5	BP3050200	60188841261	NTH_MGC	851
gb_smlt2.CB82552	-	78.50	138.70	145.85	42.9	CB82552	60188841261	NTH_MGC	851
gb_smlt2.CB83418	-	78.50	138.70	146.15	42.4	CB83418	CB83418	Rabbit corneal	424
gb_smlt2.BP027923	-	78.50	138.15	156.84	45.1	BP027923	219433	MARC 2B90	BO
gb_smlt2.AV2077923	-	78.50	137.95	160.80	46.1	AV2077923	RRC1 -33-10169	TP	461
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	851
gb_smlt2.BG649485	-	78.50	133.20	229.05	63.9	BP649485	60166091981	NTH_MGC	


```

596 AGGCGAGTTCCTCCATCCATCCATCCGACACATCCGACGGCCGAT 645
92 ATGCGAAGsergilyleuserThr...GlyTrp**GlnIuserLysIle 107
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
646 CCGCTTACTATGTATGACACACACGCGCCCTACAGACAGCATGACC 695
107 uLengIuLeuThrGlyProIyGValLeu 116
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
696 CCGTGCACCTTGCTGTCACACGACCTTA 723

seq_name: gb_ests:AA530090

seq_documentation_block:
LOCUS AA530090
DEFINITION v137c01.f1 Striatum mouse diaphragm (#937303) Mus musculus cDNA
clone IMAGE:931200 5' similar to SM:AB02.HUMAN.P04217
ALPHA-1B-GLYCOPROTEIN.; mRNA sequence.
ACCESSION AA530090
VERSION AA530090.1 GI:2272796
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 506)
Marrin M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gastelienberg, K., Stepien, M., Tan, F., Martin, J., Morris, M.,
Schellander, P., Wyler, T., Lennon, G., Soares, B., Wilson, R. and
Holeston, P.
The MASH-HMI Mouse EST Project
Unpublished (1999)
COMMENT
Contact: Marra M/Mouse EST Project
MASH-HMI Mouse EST Project
Washington University School of Medicine
660 South Euclid Avenue
Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:536120
Seq primer: -28m13 rev1 ET from Aneursham
High quality sequence stop: 287.
Location/Qualifiers
1..506
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:931200"
/clone_11b="Striatum mouse diaphragm (#937303)"
/size_type="diaphragm"
/size_note="diaphragm (kanamycin resistance)"
/note="Organ: diaphragm; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally from mRNA
prepared from diaphragm muscle. Primer: Oligo dt. Average
insert size: 1.5 kb. Uni-ZAP XR Vector: -5' adaptor
sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'
CTCATGCTTTTCTTTTCTTTT 3'
BASE COUNT
120 a 124 c 144 g 116 t
ORIGIN

```

```

10 LeuTrpGlyValThrTrpGlyProValThrValAlaIleLeuPheTyrG1 26
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
1 CTCGGGGGCTTACTGTCGCCACAGAAATACCTTAATGCTGCAATTCG 50
26 uThrGln***SerLeuTPalAcIu_SerGluHis***LeuIyThrLeu 42
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
51 CAGTAACTCAAAATGTGGGACAGACCTAGACCTGCT...GAAACCTG 97
43 GlyIleCysaspAlaIasPvalIProdyIProProGlyIasPserAenP 59
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
98 GCCAACCCTACCTGCTGTGACAGTTATGATTTGGCTCAAGAGCTTCG 147
59 Ala_ValGlnIuTrpGlyValAlaGlnIuProValHisLeuAspSerPro 75
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
148 AGCTATCTCAAAAGCGGTGTGCTGCTCAAGCAATGCTCCCTTACAG 197
76 AlaIleuHisGlnPheLeuLeu.....ThrGlyAspThrGlnG1 89
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
198 GTCTGTATACCGCTTTTCCCTGGAGGGCATTCAGTAAAGTAAACAG 247
89 YANGTATGCGAAGsergilyleuser.....ThrGlyAspThrGlnG1 98
248 CATCTACCGTCCAGATGTGGCTGACGACCCCTTCGACATTCACCTG 297
99 .....ThrGlyTrp**GlnIuserLysIleuGlnIuLeuThrGly 112
||||| 111 :||| :||| :||| :||| :||| :||| :||| :||| :|||
298 CAGCAGTCAACAGTCAACGACATCAAGCATCTGTCAGAGTCAAGG 347
113 ProIyGValLeuAlaCysSerLeuAla 121
348 AAGACCCCTTCCTGCGGCTTGCT 374

seq_name: gb_ests:BC706583

seq_documentation_block:
LOCUS BC706583p1 NIH-MGC_96 Homo sapiens cDNA clone IMAGE:4795550 5'
DEFINITION mRNA sequence.
ACCESSION BC706583
VERSION BC706583.1 GI:13982072
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Chiroptera; Homnidae; Homo.
1 (bases 1 to 531)
NIH-MGC http://mgi.cni.nih.gov/.
National Institutes of Health/ Mammalian Gene Collection (MGC)
COMMENT
Contact: Robert Strusberg, Ph.D.
Email: cga@dsi.femall.nih.gov
Funding: National Institutes of Health, M.D. Ph.D.
Genetic Library Preparation: Michael J. Brownstein (MNCRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA Sequencing by: MGC Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:
http://image.llnl.gov
1 column: 1 column: 15
High quality sequence stop: 521.
Location/Qualifiers
1..531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4795550"
/clone_1b="NIH-MGC_96"
/clone_1b_1="NIH-MGC_96"
/note="Library: Mammalian"
/nab_host="pBluescript SK"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and

```


[illegible]

```

REFERENCE      Mammalia: Eutheria; Cetartiodactyla; Suidae; Sus.
AUTHORS        1 (bases 1 to 548)
                Enhemking,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
                Scorel,Y., Watson,M.P., Grosser,W.M., Bemecl,G.A., Laegreid,W.K.,
                design and use of two pooled tissue normalized cDNA libraries for
                EST discovery in swine
TITLE          Unpublished (2000)
JOURNAL        Contact: SMIT TPL
COMMENT        PO Box 166, Clay Center, NE 68933-0166, USA
                Tel.: 402 763 4350
                Email: smith@unl.nebraska.edu, qxy
                Single pass sequencing. Bases called and alt.trimmed with phred
                v0.980904.e. Vector identified by cross_match with the -minscore 18
                and -mismatch 12 options.
FEATURES       PCR Primers
                FORWARD: AGGAGACACTATGACAT
                REVERSE: GTTCTTCACCGACGAG
                Primer 109
                Seq primer: ATTACCTGCACATCATC
                Location/Qualifiers
                1..548
                 /organism="Sus scrofa"
                 /db_xref="taxon:9833"
                 /clone_lib="MRC 1P10"
                 /library="Unpub"
                 /notes="Library from pooled tissue from day 11, 13, 15, 20,
                 and 30 embryos."
BASE COUNT    119 a      187 c      149 g      93 t
ORIGIN
alignment_scores
Quality:      92.50      Length:      81
Ratio:        1.968      Gaps:         4
Percent Similarity: 58.025   Percent Identity: 32.099
Alignment_block
US-09-471-276-831 x BG733169 ..
Align seg 1/1 to: BG733169 from: 1 to: 548
16 GTPcvtatrrtctatataatalephetygtutnchin***seantur 32
|||||.....|||
87 GCGCCAGTCCCCCACGCCCACTGGCTAC.....TGACAGGAGGAGCGCG 130
|||||.....|
32 PAlAclusecrluuhs**leauystrlneluglcncysaaaplasy 49
|||||.....|
131 CGCGCTCTCCGTGA.....Gdctcagagacccgaggaacattcttc 171
49 alPrOgiLyPro.....ProgiYapseraryLeaUPRoalaval 61
:::|||||
172 TGCCGACGACGAGACACTCTTAGAGGAGCAGACCdGd.....GAGGcg 215
62 GlnGltTrgcLYALAGInGLPrOVALHSLeLanASepPRoALAtIdly 78
:::|||||
216 AGAGAGCTGGGAGATCCCGTCCCGCCACATCCACACTGTGAAGACCCCTGTTAC 265
78 sHlsgInPhleueanthttcglayArPhgGlnLYAnqYfAYg 92
|||.....|
266 ACACCCCCTCCTCATGATAGCGAGACTGACATCCAGGAGACA 308
seq_name = gp.est2-768210
seq documentation_block:
LOCUS       T68210                      97 bp     mRNA                    EST
DEFINITION  yc40b03.t1 StrataGene Library (#937224). Homo sapiens cDNA clone
IMAGE:81117 5' similar to SP:AB210MAN F04217 ;, mRNA sequence.

```

[illegible]

LOCUS	BG619104	809 bp	mRNA	EST	22-MAY-2001
DEFINITION	607621193ef1 NCL-GCAP Brn67 Homo sapiens cDNA clone IMAGE:4932190				
ACCESSION	BG619104				
VERSION	1				
KEYWORDS	mRNA sequence.				
SOURCE	Homo sapiens				
ORGANISM	human.				
REFERENCE	Eukaryotic; Mollusca; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
TITLE	Mammalia (bases 1 to 809)				
JOURNAL	NIH-NEC http://www.ncbi.nlm.nih.gov/				
COMMENT	Unpublished NIH grant support number 1U99S				
	Contact: Robert Strassberg, Ph.D.				
	Email: rstrass@fema1.nhl.gov				
	Tissue Procurement: David N. Louis, M.D.				
	CDNA Library Preparation: Life Technologies, Inc.				
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	cDNA Sequencing by: InCyte Genomics, Inc.				
	Clone Distribution: NCI-CRCC Distribution Information can be				
	found at: http://image.llnl.gov				
	Plate: LLN410857 row: i column: 23				
	High quality sequence stop: 800.				
FEATURES					
source	Location/Qualifiers				
	1..809				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_id="BG619104"				
	/clone_1lb="NCI-GCAP_Brn67"				
	/tissue_type="templastic oligodendrogloma with Ip/19q				
	loss"				
	/lab_host="DH10B (71 phage-resistant)"				
	/note="Origin: brain; Vector: PCMV-SPORE; Site: 1; NCBI:				
	Site2: SalI; Cloned unidirectionally. Primer: Oligo dT				
	Average insert size 2.1 kb. Constructed by Life				
	Technologies, Inc. This is a NCI-GCAP Library."				
BASE COUNT	252 a 180 g 180 s 197 t				
ORIGIN					
	alignment_scores:				
	quality:	91.00	length:	75	
	ratio:	2.022	gaps:	2	
	percent similarity:	60.000	percent identity:	34.667	
alignment_block:					
OS-05-471-276-831 x BG619104/rev ..					
Align seq 1/1 to reverse of: BG619104 from: 1 to: 809					
50 ProClgPProPrGciGlyAspSerArgLeuProAlaValAlaGlnGluTrpGlyAl					
231 CAGGGCAAGAATCTGTGTGGAGAGCGTTTCCTCTCCCTCCCTGCGACTG	184				
66 acLgInUpProValAlnIsLysAspSerProAlaLeuGlnScgTrpAheLeu	83				
183 ...CATCTGAAAGAGGCGAGCACGCCATCATTTTAATAAGAGATTAACT	138				
83 eutRncLYaSpErGcInGlyArgTyArGYSArSerGLyLeuSerTrp	99				
137 TGACTCGAGGAACAAGCACTCTTA	100				
100 GYTrp**GlnLeuGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	116				
99 TGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	50				
116 uALAcYSerLeuAlaLeuArgGly	124				
49 AACAGCTTGCTGCTCCCGACAGATGCA	25				
seq_name: gb_esc2: B0775650					


```

49 valp1proci1proci1yaspser1arg1euproalavalg1ing1utrp1 65
      |||
      |||
358 AAGAGGGAGAGCGA.....GACCCCTGGCA 383
      |||
65 yalag1nc1p1roval1h1slc1u1asp1ser1proal1el1y1a1h1s1d1p1c1 82
      |||
      |||
384 CAGCAATACCCAGCTG.....GACCAAGCAAGCAAGCCGAGATCT 144
      |||
      |||
82 eule1ut1h1c1g1y1a1p1h1r1c1.....g1y1a1r1q1y1a1r1c1y1a1r1s1 95
      |||
      |||
425 TCATCTCATCTCATCAACACCACTCAACAGCAAGATACCCCTGGCACTAT 474
      |||
      |||
96 g1y1e1u1s1e1r1h1c1g1y1t1r1p1**g1n1e1u1s1e1r1y1s1e1u1d1u1e1 110
      |||
      |||
475 TACAGCTCTCAGCGCTGCTCAGACCCAGCAAGCCCTGGCACTG 519
      |||
      |||
seq_name: gB-est2:BF441074

seq_documentation_block:
LOCUS      BF441074               470 bp      mRNA
DEFINITION  257194 MARC ZPIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION  BF441074
VERSION    BF441074.1 GI:11501165
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa. Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
REFERENCE  1 (bases 1 to 470)
AUTHORS    Fabrington,S.C., Finking,B.A., Bohrer,C.A., Smith,T.P.L., Casas,E.,
          Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.W.,
          and Keele,J.W.
          Design and use of two pooled tissue normalized cDNA libraries for
          EST discovery in swine
          Mammalian Genome 12(2001)409
          Contact: Smith TPI
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smiththermal1@marc.usda.gov
          This sequence was generated and all trimmed with paired
          end sequencing. The sequence was identified by cross-match with the
          v0.960801 database. The sequence was identified by cross-match with the
          m1n1score 18
          and -m1n1match 12 options.
          PCR Primers
          FORWARD: AGCAAGACGCTATGACCAT
          BACKWARD: GTTTTCAGCTACGACGAG
          Plate: 83 row: F column: 18
          Seq primer: ATTTACGACGCTACCAATCA.
FEATURES             source
     source          1..470
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_11b="MARC ZPIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /accession="US-09-471-276-831"
                     /sequence="1"
                     /library="pooled"
                     /library_name="pooled"
                     /library_type="cDNA"
                     /library_source="placental, pituitary, and placenta."
                     /library_medium="12 g 80 T 1 others"
BASE COUNT          101 a 156 c 132 g 80 t
ORIGIN
alignement_scores:
      Quality: 87.00      Length: 54
      Ratio: 2.485
      Percent Similarity: 64.815      Percent Identity: 35.185
alignment_block:
US-09-471-276-831 x BF441074
Align seg 1/1 to: BF441074 from: 1 to: 470

```

```

43 g1y1c1i1n1c1y1a1s1p1a1s1y1a1p1roci1proci1.....Proci1y1a1s 55
      |||
      |||
71 GAGCAGGAGAGAGAGCTCTCTCCAGCCGAGCAAGACACTTACCTAGAGAGCA 120
      |||
55 g1s1e1r1a1r1e1u1p1roal1a1y1a1g1i1ng1ut1r1p1y1a1a1g1i1ng1ut1r1p1a1h1s1d1p1c1 72
      |||
      |||
121 CAGCCG.....GAGCTGACAGAGCTGGAGATACCCCTGGCACTACCAT 164
      |||
      |||
72 eule1ut1h1c1g1y1a1p1h1r1c1.....g1y1a1r1q1y1a1r1c1y1a1r1s1 88
      |||
      |||
165 GTGAGAGACACCCCTCTTTCACACAGCCGCTCTCATGATGAGGAGCA 214
      |||
      |||
89 g1y1a1r1y1a1r1g 92
      |||
      |||
215 TCCAGGAGCAGA 226
      |||
      |||
seq_name: gB-gas:TA319D090

seq_documentation_block:
LOCUS      TA319D090             556 bp      DNA
DEFINITION  TA319D090
          T-brucei sheared genomic DNA clone 319d09, reverse sequence,
          genomic survey sequence.
ACCESSION  AL492799
VERSION    AL492799.1 GI:11867569
KEYWORDS   GSS.
SOURCE     Trypanosoma brucei.
ORGANISM   Trypanosoma brucei.
Mammalia; Eukaryota; Kinetoplastida; Trypanosomatidae;
          Trypanosoma.
REFERENCE  1 (bases 1 to 556)
AUTHORS    Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
          Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
          Melville,S.E., Rajandream,M.A., and Barrell,B.G.
          Direct Submission
          Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
          project. The genome of the causative agent of human African
          trypanosomiasis, Trypanosoma brucei, was sequenced by the
          Cambridge CR10 15M, E-mail: barrell@sanger.ac.uk and
          nh@sanger.ac.uk
          Constructed at the Institute for Genomic Research (TIGR),
          Rockville, MD. Genomic DNA isolated from a cloned population of
          Trypanosoma brucei (TREG927/4 Guinea 10-1) was mechanically sheared
          to give a tight size distribution (10-11) was mechanically sheared
          to give a tight size distribution (10-11) was mechanically sheared
          described in detail in Smith, H. and Venter, J.C. (making small)
          Insert libraries for whole genome shotgun sequencing projects. In
          Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
          Barrell, Oxford University Press, 1999).
          Email: nh@sanger.ac.uk
          Details of T. brucei sequencing at the Sanger Centre are available
          at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             source
     source          1..556
                     /organism="Trypanosoma brucei"
                     /strain="TREG927"
                     /db_xref="taxon:5691"
                     /clone="319d09"
BASE COUNT          124 a 157 c 147 g 128 t
ORIGIN
alignement_scores:
      Quality: 87.00      Length: 110
      Ratio: 1.450
      Percent Similarity: 54.545      Percent Identity: 30.000
alignment_block:
US-09-471-276-831 x TA319D090
Align seg 1/1 to: TA319D090 from: 1 to: 556
      31 leu1r1p1a1a1g1a1.....SeC1u1h1s1**Leu1y1s1h 41
      |||
      |||

```

```

183 ATTGGGAAAGCTTCAGACTGCTCCCGACACATCTCCGAAGC 232
41 tleuGIyGInCysaspIaAspvalProdiTyProProdiAspserArgL 58
   ::::: ||||| :::: |||
233 CCGCACTCCACAGACACAGACGTCCTCTCCCGGCGGCGCCGACGACA 282
58 euProAlaValGInGInuTrIPdIyAla.....GInGInuProValHisLen 72
   ||||| :::: |||||
283 GTGACCGCGCTCACCGCAGCGACGACAGAAAGACCTTCCAGCTC 332
73 AspserProAlaIleIySHISGInPhe.....LeuLeuThrgIyAspPh 87
   ||||| :::: |||||
333 CCGCTCCCGCAAGATACACATCCCTCCGCGCTTCGATCGACGAATA 382
87 tGInGInuArgTyArgCysArgSerGlyeuserThrdIyTrp***GInL 104
   ::::: ||||| ::::: |||||
383 TAAAGGGAGATTACGTTCGACAAACCCCTGCATGAAAGCGCGCTGTC 432
104 euserIyLeuGInGInuLeuThrgIyPro.....Iys 114
   ||||| :::: |||||
433 TAGAGACGATTTAGATTGCTTACCTCTCCACAGCTCCGACAGCGTCC 482
115 ValLeuAlaIcySerIeuAlaIeuAspGly 124
   ||||| :::: |||||
483 GTCCGCGCGCTCTCCACGCGCTGACCGG 512

seq_name: gb_gas:A0939939
seq_documentation_block: 598 bp DNA GSS 27-JAN-2001
LOCUS BC382087
DEFINITION Sheared DNA-43C17, RF, Sheared DNA Trypanosoma brucei genomic clone
VERSION A0939939
GSS:
KEYWORDS Trypanosoma brucei.
SOURCE Eukaryote: Euglenozoa: Kinetoplastida: Trypanosomatidae;
ORGANISM Trypanosoma.
1 (bases 1 to 598)
E1-Sayed,N., Zhao,H., Gill,S., Sub,E., Malek,J., Fujii,C.,
Gerard,C., Leech,V., de Jong,P., Olin,E., Melville,S., Donelson,J.,
Fraser,C. and Adams,M.
Determination of clone end sequences from Trypanosoma brucei GUT4
Impublished (1999)
Other GSS: Sheared DNA-43C17, TR
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0201
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUT4 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tldb/tldb/.
Seq primer: M13-forward
Class: shocunt.
ORIGIN
1 598
//
/organism="Trypanosoma brucei"
/strain="TREU927/4 GUT4 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-43C17"
/clone_lib="Sheared DNA-43C17"
/contig="1"
/contig_lib="ProdiTyProProdiAspserArgL"
/institute="Genomic Research (TIGR), Rockville, MD."
/genome="Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUT4 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is

```

```

BASE COUNT      136 a      165 g      174 g      123 t
ORIGIN

alignment_scores:
Quality:      87.00      Length:      110
Ratio:        1.450      Gaps:         4
Percent Similarity: 54.545      Percent Identity: 30.000

alignment_block:
US-09-471-276-831 x A0939939

Align seg 1/1 to: A0939939 from: 1 to: 598

31 LeuTrpAlaGlu.....SerGInHis***LeuIyTrh 41
   ::::: ||||| ::::: |||||
108 ATTGGGAAAGCTTCAGACTGCTCCCGACACATCTCCGAAGC 157
41 tleuGIyGInCysaspIaAspvalProdiTyProProdiAspserArgL 58
   ::::: ||||| ::::: |||
158 CCGCACTCCACAGACACAGACGTCCTCTCCCGGCGGCGCCGACGACA 207
58 euProAlaValGInGInuTrIPdIyAla.....GInGInuProValHisLen 72
   ||||| :::: |||||
208 GTGACCGCGCTCACCGCAGCGACGACAGAAAGACCTTCCAGCTC 257
73 AspserProAlaIleIySHISGInPhe.....LeuLeuThrgIyAspPh 87
   ||||| :::: |||||
258 CCGCTCCCGCAAGATACACATCCCTCCGCGCTTCGATCGACGAATA 307
87 tGInGInuArgTyArgCysArgSerGlyeuserThrdIyTrp***GInL 104
308 TAGAGACGATTTAGATTGCTTACCTCTCCACAGCTCCGACAGCGTCC 357
104 euserIyLeuGInGInuLeuThrgIyPro.....Iys 114
   ||||| :::: |||||
358 TAAAGGGAGATTACGTTCGACAAACCCCTGCATGAAAGCGCGCTGTC 407
115 ValLeuAlaIcySerIeuAlaIeuAspGly 124
   ||||| :::: |||||
408 GTCCGCGCGCTCTCCACGCGCTGACCGG 437

seq_name: gb_ast2:BC382087
seq_documentation_block: 506 bp mRNA EST 12-MAR-2001
LOCUS BC382087
DEFINITION 297853 MARC JPLG Sus scrofa CDNA 5', mRNA sequence.
VERSION BC382087.1 GI:13306559
KEYWORDS EST.
SOURCE Sus scrofa.
ORGANISM Sus scrofa
Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Cetartiodactyla: Suidae: Sus.
1 (bases 1 to 506)
Fahnestock, G., Winking, B.A., Bohrer, G.A., Smith, T.P.L., Coase, E.,
Stoner, R.T., Heaton, W.P., Grose, M.M., Bennett, G.A., Langfield, W.M.,
and Keefe, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TP.
ORIGIN
1 506
//
/organism="Sus scrofa"
/strain="MARC JPLG Sus scrofa CDNA 5'"
/db_xref="taxon:10090"
/clone="MARC JPLG Sus scrofa CDNA 5'"
/clone_lib="MARC JPLG Sus scrofa CDNA 5'"
/institute="US Meat Animal Research Center,
National Animal Health Research Center,
P.O. Box 4706, Ames, IA 50013, USA"
/tel="515-281-4366"
/fax="515-281-4366"
/email="fahnestock@iastate.edu"

```

v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
 FORWARD: AGCAACACATCATACCAT
 REVERSE: GTTTTCCATCATACGACG
 Plate: 96 row: 12
 Seq primer: ATTATGCTGACATCATACG

FEATURES
 source
 1..506
 /organism: "Sus scrofa"
 /db_xref: "taxon:9823"
 /clone_lib: "MARC 1P1G"
 /library: "pooled"
 /lab_host: "DH10B"
 /note: "Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT
 93 a 131 c 177 g 105 t

ORIGIN

alignment_scores:
 Quality: 86.50 Length: 74
 Ratio: 2.012 Gaps: 4
 Percent Similarity: 58.108 Percent Identity: 32.432

alignment_block:
 US-09-471-276-831 x BG382087/rev ..

Align seg 1/1 to reverse of: BG382087 from: 1 to: 506

```

16 GCTGTCGATGThGtAlaAlaTlePheryTgTlThGIn***SeTcentTt 32
|||||
205 GGGCCAGTCCGCCACCCCTCCATGCTAC.....TGACGAGAGAGCGTG 162
|||||
32 pAlaGtUsScGtUHLs***LeuLygThTleGtGtGtGtGtGtGtGtGt 49
|||||
161 GGGCTTCCTCGAA.....GGTCTAGAGCGCGAGAGACCTCTCTC 121
|||||
49 aAlPrGtUPro.....ProGtYAspScTgTleUpProAlaVal 61
|||||
120 TCCCCAGCGCGAGACACTCTGAAAGCGCCGAGCGCG.....GAGGTG 77
|||||
62 GtGtUtrPcTlYAlaGtInUpProVAHtSLeUAspScTgProAlaTleLy 78
|||||
76 AGGAGCTCGGAGTCCCTCCGCCATCATCTGACGTGAGACCTCGTTTAC 27
|||||
78 sHtISgInPhLeuLeuThrGly 85
|||||
26 ACACGCCCTCTCTCATGATGCG 5

```

seq_name: gb_cst2:BG609827

seq_documentation_block:
 LOCUS BG609827 548 bp mRNA EST 17-APR-2001
 DEFINITION 323811 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BG609827
 VERSION BG609827.1 GI:13659806
 KEYWORDS EST.

ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 548)
 Fahnenstang,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Ganss,E.,
 and Kessler,W. National.M.F., Grosser,W.H., Benneke,G.A., Laegreid,W.W.,
 and Kessler,W. National.M.F., Grosser,W.H., Benneke,G.A., Laegreid,W.W.,
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 TITLE
 National Institutes of Health
 COMMENT
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smilhemall@marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
 FORWARD: AGCAACACATCATACCAT
 REVERSE: GTTTTCCATCATACGACG
 Plate: 96 row: 12
 Seq primer: ATTATGCTGACATCATACG

FEATURES
 source
 1..548bp "Sus scrofa"
 /db_xref: "taxon:9823"
 /clone_lib: "MARC 1P1G"
 /library: "pooled"
 /lab_host: "DH10B"
 /note: "Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT
 113 a 188 c 145 g 102 t

ORIGIN

alignment_scores:
 Quality: 86.50 Length: 74
 Ratio: 2.012 Gaps: 4
 Percent Similarity: 58.108 Percent Identity: 32.432

alignment_block:
 US-09-471-276-831 x BG609827 ..

Align seg 1/1 to: BG609827 from: 1 to: 548

```

16 GCTGTCGATGThGtAlaAlaTlePheryTgTlThGIn***SeTcentTt 32
|||||
342 GGGCCAGTCCGCCACCCCTCCATGCTAC.....TGACGAGAGAGCGTG 385
|||||
32 pAlaGtUsScGtUHLs***LeuLygThTleGtGtGtGtGtGtGtGtGt 49
|||||
386 GGGCTTCCTCGAA.....GGTCTAGAGCGCGAGAGACCTCTCTC 426
|||||
49 aAlPrGtUPro.....ProGtYAspScTgTleUpProAlaVal 61
|||||
427 TCCCCAGCGCGAGACACTCTGAAAGCGCCGAGCGCG.....GAGGTG 470
|||||
62 GtGtUtrPcTlYAlaGtInUpProVAHtSLeUAspScTgProAlaTleLy 78
|||||
471 ACACGCCCTCTCTCATGATGCG 520
|||||
78 sHtISgInPhLeuLeuThrGly 85
|||||
521 ACACGCCCTCTCTCATGATGCG 542

```

seq_name: gb_cst2:BG330746

seq_documentation_block:
 LOCUS BG330746 874 bp mRNA EST 27-FEB-2001
 DEFINITION 602430265F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4547592 5', mRNA sequence.
 ACCESSION BG330746
 VERSION BG330746.1 GI:1317184
 KEYWORDS EST.

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 874)
 National Institutes of Health
 COMMENT
 Contact: Robert Strausberg, Ph.D.


```

TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Stranberg, Ph.D.
E-mail: rstranber@nhi.nih.gov
Title: Hsp70 gene family
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Inyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Accession Number: U08691
High quality sequence stop: 726.
FEATURES
source
1..945
/organism='Homo sapiens'
/db_xref='taxon:9606'
/clone_1 IMAGE:453199*
/clone_2 Lib:NIH_MCC_46*
/lab host='DH10B (phage-resistant)'
/vector='Organ: uterus; Vector: pGB7'
/notes='cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCG(G). Size-selected >500bp for average insert size
1.kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
and submitted to the National Cancer Institute (NCI)
Library of Human Genome Resources (LHGR) as NIH_MCC
II R1 (Life Technologies). Note: this is a NIH_MCC
library.'
BASE COUNT
216 a 219 c 285 g 225 t
ORIGIN
alignment_scores:
Quality: 85.00 Length: 69
Ratio: 1.771 Gaps: 3
Percent Similarity: 53.933 Percent Identity: 31.461
alignment_block:
US-09-471-276-831 x BC337700/rev ..
Align seq 1/1 to reverse of: BC337700 from: 1 to: 945
41 Thrleudlgicndyapmlasvpaalpavprociyproyjaysesrvc 57
||||| :||| ||||| |||||
774 ACACGTGACTCTCACACATGCCCAACAGAGCGCTCCACCCTCCACGAC 75
57 gteppolaavlacigntatctttgsgstgcgcccccacatcatcgtac 75
724 TGTCCAGCATGACGATCTCTTTCGSGSTGCGCCCCACATCATCATC 675
72 LeuapserpolalilelyshllaglnhelmauleuthriclYAspThcI 88
||||| :||| ||||| |||||
674 CAAGCTTCGCCACGCCAGCACACGAC.....ACCCA 643
88 nglYahgrtyrAryCYsarGserGylseurThrGlYTP**GInLeus 105
642 ACCCGCGCTTTCTTGACAGCTCAGAGCTTTGGCATCTCTTCGCTCGA 953
105 acTyaseuleuclguethrlcgLYproLyysVlaeulaAcyssectueAla 121
||||| :||| ||||| |||||
592 CTGACACGTGACATCATGACAGATCAATTCAATCTCTTCATAAACGC 543
122 Leuapcylyalaser 126
||||| :||| ||||| |||||
542 CTTAAAGGTGCGCAC 528
seq_name: gb_est12:BC336925
seq_documentation_block:
LOCUS BC336925 886 bp mRNA EST 12-MAR-2001
DEFINITION BC243382991 NIH_MCC_20 Homo sapiens cDNA clone image:4552014 5'.

```

```

ACCESSION      mRNA sequence.
BC356925
VERSION        BC356925.1
KEYWORDS       EST.
SOURCE         Human
ORGANISM       Homo sapiens
REFERENCE      Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi:
AUTHORS        Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
TITLE          NIH-MGC project: NIH.gov/.
MISCINFO       National Institutes of Health. (MGC)
COMMENTS       (bases 1 to 886)
                Published 1993. Transcribed. Ph.D.
                Email: genbank@ncbi.nlm.nih.gov
                Tissue procurement: ATCC/PCRD/DRP
                CDNA Library Preparation: Ling Hong/Rubin Laboratory
                CDNA Library Arrived by: The I.M.A.G.E. Consortium (LLM).
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLM at:
                http://www.genome.gov/
                Plasmid: LHM247 rev: 1 column: 07
                High quality sequence spot: 702.
                Location/Qualifiers
                1..886
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /gene="MMS19"
                /chromosome="15"
                /feature="exon"
                /tissue_type="melanotic melanoma"
                /note="Organ: skin. Vector: pOT79; Site.1: XhoI; Site.2:
                EcoRI; CDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adapter: GGCGAGCTG. Size-selected 350bp for average
                length. Library prepared by the laboratory of Gerald M. Rubin (University of Hong In
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT     215 a 203 c 263 g 205 t
ORIGIN
alignment_scores:
      Quality:      84.50      Length:      80
      Ratio:        1.837      Gaps:        4
      Percent Similarity: 57.500      Percent Identity: 32.500
alignment_block:
  05-09-471-276-631 x BC356925/rev ..
Align seg 1/1 to reverse of: BC356925 from: 1 to: 886
50 ProGIGProProGIGYspASpser...LeuPRAIaVIdIGIurPdI 65
|||||
746 CCAAGCGCCCTCCACAGCCACAGCAAGTGTACCACACAGCGATGTGGCG 697
65 yAla..GInGluPrcVAlHIsLeuAspSPePPoAlaIEtIElySHtGInp 81
696 TGGCCCCCAGCGATGATGACGAGATGTCCGACGACAGCACACACAC 648
81 heIeuIeuthrGtIAspHrGInGlyArgTyrTArGcYsArgSercIy... 96
647 .....ACCGAGACCCGCGCTTTCACGCGCTGGCTTGG 615
97 LeuSerThrGlyTrp**GInLeuSerTyrIleuLeuGluIenPhrGlyPr 113
|||||
614 CTTCGCACTTCCTTGCGTCACTGTCAAGCTCACTGACATCAACAGAGA 565
|||||
564 TCAAAATCATATGCTGTGCACAAAGCGCTTAAAGTCCCAAT 525

```



```

seq_name: gb_est1:BE542780
seq_documentation_block: 1025 bp mRNA EST 09-AUG-2000
LOCUS BE542780.1 MIM_MGC_10 Homo sapiens cDNA clone IMAGE:345261.5',
DEFINITION mRNA sequence.
ACCESSION BE542780
VERSION BE542780.1 GI:9771425
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1025)
AUTHORS NIH-MGC http://mim.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: ATCC. Life Technologies, Inc.
CDNA Library Arrayed by: Invitae Genomics, Inc.
CDNA Sequencing by: Invitae Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
http://image.llnl.gov
Plate: LAM8434 row: 3 column: 22
High quality sequence stop: 130.
FEATURES
Location/Qualifiers
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:345261"
/clone_lib="MIM_MGC_10"
/cell_line="MGC36"
/lab_host="DH108"
/vector="pGEMT-oriX; Vector: PCW-SPOOF6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 295 a 219 c 294 g 217 t
ORIGIN
alignment_scores:
Quality: 84.00 Length: 98
Ratio: 1.750 Gaps: 4
Percent Similarity: 48.980 Percent Identity: 28.571
alignment_block:
US-09-471-276-831 x BE542780 ..
Align seg 1/1 to: BE542780 from: 1 to: 1025
10 LeuTRPGlyValThrTPGlyProValThrGluAlaAlaIlePheTy-Gl 26
||||| |||||||
457 CTGTGGCATTCGATTGGGGGGGCGGCGCTTCAAAAAGCGCTTTTTCGAGA 506
|||||
26 uThrGtI+***seTLeuTPRAlaIleuSerGluHis+***LeuThrLeuG 43
||||| |||||||
507 AGCGGCGCTTCCTTGGGGGACAA.....ATATTGAACCCCA. 546
|||||
43 LysGInGysAspAlaAspValProGlyProPro..... 53
|||||
547 ..ACGThGAGGAGCGCTGTGCTCCGCCCAACCTGCTGGGGGTATGCT 554
|||||
54 .....GlyAspSerAlaLeuPr 59
|||||
595 TTTCACAAACCCCTGTGGGGGCGCTGCTGCCACACTTGAAGAAGAAC 644
|||||
59 oAlaValoGInGlyTPRAlaGInGlyProValHisIleu.....A 73
||||| |||||||
645 TACTGTGATTAAGATTGGGGGAGAGAGTCCCTCCCAACCTGGGAGAGGCA 694
|||||

```

```

73 spSerProAlaIleuTyrHisGlnPheLeuLeuThrGlyAspThr 87
||||| |||||||
695 ACACACCCCTCTGTCAAGATATATGAGGGGACAGCCGCAACA 738
|||||
seq_name: gb_est1:AM001162
seq_documentation_block: 481 bp mRNA EST 27-OCT-1999
LOCUS AM001162
DEFINITION wu25c03.x1 Soares_Diackgrae.colon_NKCD Homo sapiens cDNA clone
IMAGE:2521060.3', mRNA sequence.
ACCESSION AM001162
VERSION AM001162.1 GI:5848078
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butirata; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 481)
AUTHORS NIH-CCAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
This clone is available royalty-free through LBLT; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: 7400P from cDNA 462.
High qual location/Qualifiers
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2521060"
/clone_lib="Soares_Diackgrae.colon_NKCD"
/tissue="type=colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH108 (phage-resistant)"
/notes="Organ: colon; Vector: pTR73D-Pac (Pharmacia) with a
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATGTGAATGGAGGCGGCGGCGCTGTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adapter
primers and then ligated to the modified pTR73 vector. Library
and Eco RI sites were the modified pTR73 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Diackgrae (Washington University,
St. Louis, MO). Samples were from moderate to severe Crohn's
disease. Samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Souza and M. 127 g 124 t.
BASE COUNT 102 a 128 c 127 g 124 t
ORIGIN
alignment_scores:
Quality: 83.50 Length: 112
Ratio: 1.750 Gaps: 2
Percent Similarity: 43.750 Percent Identity: 27.673
alignment_block:
US-09-471-276-831 x AM001162 ..
Align seg 1/1 to: AM001162 from: 1 to: 481
22 AlaIlePheTyGluThrGtI+***seTLeuTPRAlaIleuSer..... 35
||||| |||||||
37 GCACAAATTTTACGAAAGTCACATTACCCCTGACGCCAAGACCTCTTT 86
|||||
36 .....GluHis 37
|||||
87 CTCACGGGAGAGCTGTGTCACAGTGGCTGTGGTGGAGGTGATCATCAT 136
|||||

```


KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
TITLE NIH-MGC clone distribution information can be
JOURNAL found through the I.M.A.G.E. Consortium/LML at:
COMMENT http://image.llnl.gov
place: LLM411332 row: b column: 06
High quality sequence stop: 666.
location/Qualifiers
1. 919
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5135141"
/clone_lib="NCI_COAP_L19"
/vectors="DH10B (T1 phage-resistant)"
/notes="Organ library: PCW-SHORT6. Site 1: NotI;
Site 2: SalI. Cloned unidirectionally. primer: oligo dr.
Average insert size 1.9 kb. Constructed by life
Technologies. Note: this is a NCI_COAP library."
BASE COUNT 215 a 245 c 279 g 180 t
ORIGIN
alignment_scores:
Quality: 82.00 Length: 86
Ratio: 1.907 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 30.233
alignment_block:
US-09-471-276-831 x B1331839/rev ..
Align seg 1/1 to reverse of: B1331839 from: 1 to: 919
43 G1G1NCyGcyspAlaASP.....ValProG1 51
|||||
51 yProProG1yAspserATgLeuProAlaValG1nG1UTRg1yAlaG1nG 68
867 GCACACACGCTTGGCCCTGGCGTGGC...GACACATGCGATCGACAC 821
68 lUpProValHisLeuAspserProAlaLe1yHisG1nPhLeuLeuPh 84
820 CGCCTTTACCATCCACAGCCCTGGA.....ACCTGGCG 785
85 G1yAspThrG1nG1yATgATgATgATgATgATgATgATgATgATgATg 100
785 GCACACACGCTTGGCCCTGGCGTGGC...GACACATGCGATCGACAC 736
101 ...Trp***G1nLeuSer1yLeuLeuG1nG1yG1nG1yG1nG1yG1n 116
735 AACCTGGCGACAGCCGATGCTGCGACGCTGATGCTGCGACGCTGATA 685
116 eAlaLeGys 118
685 CACGATGC 678
seq_name: gb_ace2:BF347835
seq_documentation_block:

LOCUS BF347835 1101 bp mRNA EST 22-NOV-2000
DEFINITION 6020229272F1 NCI_COAP_Brm7 Homo sapiens CDNA clone IMAGE:4158569
ACCESSION BF347835
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
TITLE NIH-MGC clone distribution information can be
JOURNAL found through the I.M.A.G.E. Consortium/LML at:
COMMENT http://image.llnl.gov
place: LLM49434 row: o column: 18
High quality sequence stop: 573.
location/Qualifiers
1. 1101
/organism="Homo sapiens"
/strain="Bm7"
/db_xref="taxon:9606"
/clone="IMAGE:4158569"
/clone_lib="NCI_COAP_Brm7"
/vectors="DH10B (T1 phage-resistant)"
/notes="Organ library: PCW-SHORT6. Site 1: NotI;
Site 2: SalI. Cloned unidirectionally. primer: oligo dr.
Average insert size 1.9 kb. Constructed by life
Technologies. Note: this is a NCI_COAP library."
BASE COUNT 369 a 292 c 224 t
ORIGIN
alignment_scores:
Quality: 83.00 Length: 135
Ratio: 1.306 Gaps: 0
Percent Similarity: 50.000 Percent Identity: 27.206
alignment_block:
US-09-471-276-831 x BF347835/rev ..
Align seg 1/1 to reverse of: BF347835 from: 1 to: 1101
13 ValThrProG1yProValHisLeuAlaValG1nG1UTRg1yAlaG1nG 29
569 TTGACCTGGCGACAGCTGCGACGATGCTGCTGATATATAGCTGCT 520
29 *SerLeu.....TrpAlaG1ySerG1nHis**Leu1yG1n 42
519 GCCTGTTAAATACCTAGTGGCTTTTACAGTACGACGACGAGGACCTGT 470
42 eAlaG1n.....CysAspAlaAspValProG1yPro... 52
469 TGGACGAGCTGCGACAGAGCTGCTCCGAGTGAAGAGCTGCTGCTAGCT 420
53ProG1yAspserATgLeuProAlaValG1nG1UTRg1yAla 66
419 GCACCTGCTGACGAGCTGACAGAGCTGCTCCCTGCTGACAGCTGCTG 370
66 lAsG1nUpProValHisLeuAspserProAlaLe1yHisG1nPhLeu 82
369 CTTCGACACGCGACCTTACCTGCTGACGACGATGATGCTGCTGCGG 320
83 LeuThrG1yAspThrG1nG1yATgATgATgATgATgATgATgATgATg 92

[illegible][illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using SW model

Run on: January 7, 2002, 16:49:15 : Search time 31.99 seconds
(without alignments)

37.048 Million cell updates/sec

Title: US-09-471-276-831_COPY_1_16

Perfect score: 1 NSMLVIFLLNGYING 16

Sequence: BUCS062

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A: Geneseq_1101.*
1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	87	100.0	126	21	AA164760	Human 5' EST relat
2	51	58.6	512	22	AA187724	Human liver growth
3	48	55.2	33	22	AA186894	Human M019 signal
4	48	55.2	85	22	AA186893	Human M019 protein
5	48	55.2	85	22	AA186893	Human M019 protein
6	44	50.6	168	22	AA182182	Mouse liver growth
7	44	50.6	168	22	AA182182	Human polyprotein
8	44	50.6	200	21	AA182180	Arbidopsis thailia
9	44	50.6	228	21	AA182189	Arbidopsis thailia
10	44	50.6	280	20	AA173399	Amino acid sequenc
11	44	50.6	280	21	AA181504	Amino acid sequenc

12	44	50.6	280	21	AA186815	DNA encoding a hum
13	44	50.6	280	21	AA192504	Human OXRE-1. Hom
14	44	50.6	280	22	AA180021	Human polypeptide
15	44	50.6	280	22	AA188357	Human membrane or
16	44	50.6	280	22	AA180267	Human PRO266 prote
17	44	50.6	280	22	AA182186	Human secreted pro
18	44	50.6	280	22	AA182186	Human secreted pro
19	44	50.6	372	21	AA182363	Human secreted pro
20	44	50.6	439	21	AA182363	Human secreted pro
21	44	50.6	439	21	AA182363	Human secreted pro
22	44	50.6	541	22	AA181431	Human novel protei
23	44	50.6	541	22	AA181431	Human secreted pro
24	44	50.6	541	22	AA181431	Human secreted pro
25	44	50.6	541	22	AA181431	Human secreted pro
26	44	50.6	541	22	AA181431	Human secreted pro
27	44	50.6	541	22	AA181431	Human secreted pro
28	44	50.6	541	22	AA181431	Human secreted pro
29	44	50.6	541	22	AA181431	Human secreted pro
30	44	50.6	541	22	AA181431	Human secreted pro
31	44	50.6	541	22	AA181431	Human secreted pro
32	44	50.6	541	22	AA181431	Human secreted pro
33	44	50.6	541	22	AA181431	Human secreted pro
34	44	50.6	541	22	AA181431	Human secreted pro
35	44	50.6	541	22	AA181431	Human secreted pro
36	44	50.6	541	22	AA181431	Human secreted pro
37	44	50.6	541	22	AA181431	Human secreted pro
38	44	50.6	541	22	AA181431	Human secreted pro
39	44	50.6	541	22	AA181431	Human secreted pro
40	44	50.6	541	22	AA181431	Human secreted pro
41	44	50.6	541	22	AA181431	Human secreted pro
42	44	50.6	541	22	AA181431	Human secreted pro
43	44	50.6	541	22	AA181431	Human secreted pro
44	44	50.6	541	22	AA181431	Human secreted pro
45	44	50.6	541	22	AA181431	Human secreted pro
46	44	50.6	541	22	AA181431	Human secreted pro
47	44	50.6	541	22	AA181431	Human secreted pro
48	44	50.6	541	22	AA181431	Human secreted pro
49	44	50.6	541	22	AA181431	Human secreted pro
50	44	50.6	541	22	AA181431	Human secreted pro
51	44	50.6	541	22	AA181431	Human secreted pro
52	44	50.6	541	22	AA181431	Human secreted pro
53	44	50.6	541	22	AA181431	Human secreted pro
54	44	50.6	541	22	AA181431	Human secreted pro
55	44	50.6	541	22	AA181431	Human secreted pro
56	44	50.6	541	22	AA181431	Human secreted pro
57	44	50.6	541	22	AA181431	Human secreted pro
58	44	50.6	541	22	AA181431	Human secreted pro
59	44	50.6	541	22	AA181431	Human secreted pro
60	44	50.6	541	22	AA181431	Human secreted pro
61	44	50.6	541	22	AA181431	Human secreted pro
62	44	50.6	541	22	AA181431	Human secreted pro
63	44	50.6	541	22	AA181431	Human secreted pro
64	44	50.6	541	22	AA181431	Human secreted pro
65	44	50.6	541	22	AA181431	Human secreted pro
66	44	50.6	541	22	AA181431	Human secreted pro
67	44	50.6	541	22	AA181431	Human secreted pro
68	44	50.6	541	22	AA181431	Human secreted pro
69	44	50.6	541	22	AA181431	Human secreted pro
70	44	50.6	541	22	AA181431	Human secreted pro
71	44	50.6	541	22	AA181431	Human secreted pro
72	44	50.6	541	22	AA181431	Human secreted pro
73	44	50.6	541	22	AA181431	Human secreted pro
74	44	50.6	541	22	AA181431	Human secreted pro
75	44	50.6	541	22	AA181431	Human secreted pro
76	44	50.6	541	22	AA181431	Human secreted pro
77	44	50.6	541	22	AA181431	Human secreted pro
78	44	50.6	541	22	AA181431	Human secreted pro
79	44	50.6	541	22	AA181431	Human secreted pro
80	44	50.6	541	22	AA181431	Human secreted pro
81	44	50.6	541	22	AA181431	Human secreted pro
82	44	50.6	541	22	AA181431	Human secreted pro
83	44	50.6	541	22	AA181431	Human secreted pro
84	44	50.6	541	22	AA181431	Human secreted pro

ALIGNMENTS

```

85      39      44.8      301  21      AAV57945      Human transmembran
86      39      44.8      301  22      AAV5949      Human polypeptide
87      39      44.8      301  22      AAU1254      Human PEO343 poly
88      39      44.8      301  22      AAU1254      Human PEO343 poly
89      39      44.8      332  21      AAV82703      Tick derived cycle
90      39      44.8      344  18      AAV26767      Human chemokine re
91      39      44.8      344  19      AAV23957      Amino acid sequenc
92      39      44.8      356  19      AAV48087      Human macrophage/d
93      39      44.8      353  20      AAV9454      G-protein coupled
94      39      44.8      353  22      AAU0479      Human G-protein co
95      39      44.8      387  21      AAV5072      Human G-protein co
96      39      44.8      387  21      AAV5072      G protein-coupled
97      39      44.8      457  22      AAU03813      S. epidermidis ope
98      39      44.8      473  22      AAV81483      C glutamicum prote
99      39      44.8      497  22      AAV90137      C glutamicum prote
100     39      44.8      499  22      AAV90588      C glutamicum prote

```

```

RESULT 1
AAV64670 ID AAV64670 standard; Protein: 126 AA.

```

```

AAV64670:

```

```

01-FEB-2000 (first entry)

```

```

Human 5' EST related polypeptide SEQ ID NO:831.

```

```

Human: 5' EST; expressed sequence tag; secreted protein; diagnosis;
gene therapy; chromosome mapping; upstream regulatory sequence;
locus; location; development; protein synthesis; stability;
regulation; identification.

```

```

Homo sapiens.

```

```

W0953051-A2.

```

```

21-OCT-1999.

```

```

09-APR-1999: 99MO-1E00712.

```

```

09-APR-1998: 98US-0057719.

```

```

28-APR-1998: 98US-0069047.

```

```

(GRST ) GRNST.

```

```

Dumas MLine Edwards J, Duclert A, Giordano J;

```

```

WPI: 2000-038446/03.

```

```

N-PSDB: AA42284.

```

```

Novel secreted protein 5' expressed sequence tag sequences used in
diagnostic, forensic, gene therapy, and chromosome mapping procedures

```

```

Claim 2: Page 603; 837pp; English.

```

AA42265 to AA43075 represent novel 5' expressed sequence tag (EST) sequences, corresponding to human secreted proteins. AAV64651 to CC AAV6438 represent the EST-related proteins corresponding to AA42265 to CC AA43072. The 5' ESTs can be used for producing secreted human gene products. They can be used to identify and isolate 5' untranslated regions (UTRs) and upstream regulatory regions which control the expression of genes. The ESTs are also useful as probes for CC well as stability of mRNA. The ESTs are also useful as probes for CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can CC also be used in forensic procedures to identify individuals, or in CC diagnostic procedures to identify individuals having genetic diseases CC resulting from abnormal gene expression. The products may also be used in CC gene therapy protocols. The nucleic acids encoding signal peptides can be used for directing extracellular secretion of a polypeptide or the

insertion of a polypeptide into a membrane, or importing a polypeptide into a cell. The proteins encoded by the EST sequences may be useful in treating a variety of human conditions. Secreted proteins have value in forensic procedures. AA42265 to AA43072 and AAV6464 to AAV64650 represent sequences used in the exemplification of the present invention.

Sequence 126 AA:

```

Query Match 100.0% Score 87; DB 21; Length 126;
Expect 1.0e-08; Positives 10; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSMLVFELLNGVTNC 16
1 msmlvfeLLlwgvtwg 16

```

```

RESULT 2

```

```

AA48724 ID AA48724 standard; Protein: 512 AA.

```

```

AA48724:

```

```

09-MAR-2001 (first entry)

```

```

Mouse liver growth hormone-induced clone 5 ORF #1, SEQ ID NO:8.

```

```

Mouse: growth hormone; GH regulatable gene; liver pathology; hypertrophy;
hepatocellular lesion; hyperplasia; altered expression level; clone 5;
diagnostic marker; gigantism; acromegaly; diabetes; hepatotropic;
transgenic animal; drug screening; drug discovery; murine;
cytoplasmic protein; open reading frame; ORF.

```

```

Mus sp.

```

```

W0200066787-A2.

```

```

09-NOV-2000.

```

```

05-MAY-2000: 2000MO-US12366.

```

```

05-MAY-1999: 99US-0132663.

```

```

(UYOH-) UNIV OHIO.

```

```

Kopchick JJ, Tjong J;

```

```

WPI: 2001-007239/01.

```

```

N-PSDB: AAC87257.

```

Diagnosing abnormal levels of growth hormone activity in liver CC comprising assaying growth transcriptional activity and protein CC expression level of hormone-regulatable liver genes, as diagnostic CC markers of liver pathology -

Claim 2: Page 53; 65pp; English.

The invention relates to a method of diagnosing abnormal levels of growth hormone (GH) activity in the liver, or predicting a change in the condition of the liver in response to abnormal GH activity. The method involves correlating the level of expression of certain specific genes with the level of GH activity in the liver, or with an expected change in the condition of the liver as the result of GH activity. Excessive GH increases liver size as a consequence of both hyperplasia and hepatocyte hypertrophy, and hepatocellular lesions which progress with age. Studies in transgenic mice which express high levels of bovine growth hormone identified a number of genes whose expression in the liver is altered by high GH levels. The genes which are upregulated are those encoding alpha-fetoprotein, corticosteroid binding globulin, retin, table-interacting protein, paraoxonase-3, cytochrome P45011A,

CC 5-2 hydroxyacid oxidase, interferon alpha/beta receptor, growth hormone
 CC receptor, proteasome 2-subunit, and coagulation factor V. Two novel
 CC genes, clone 5 (AAC87257) and clone 45 (AAC87258) are also upregulated
 CC in response to abnormally high GH levels. Conversely, expression of the
 CC gene encoding 3-beta hydroxysteroid dehydrogenase/delta-5-delta-4
 CC isotype 1 protein is downregulated in a GH-inhibition rat model related to
 CC transgenic human animals expressing a GH-inhibition transgene. These
 CC exhibit or have a propensity to develop a liver pathology, an assay for
 CC drugs which inhibit the development of, or which treat a liver pathology,
 CC comprising administering the drug to the transgenic animal; and
 CC preventing or treating a liver pathology in a patient comprising
 CC administering a drug which inhibits the expression of a GH-induced gene.
 CC The method of the invention is used for diagnosing abnormal levels of GH
 CC in patients with abnormal levels of GH activity. The GH-regulatable
 CC liver genes and proteins are useful as diagnostic markers of liver
 CC pathology. Assays for the expression of these genes is useful for the
 CC diagnosis of liver pathologies associated with gigantism or acromegaly or
 CC with diabetes, as other causative agents may act directly or indirectly
 CC upon the same genes. The present sequence represents the protein encoded
 CC by the gene encoding the novel mouse liver cDNA clone 5.
 CC The protein is thought to be cytoplasmic.

SO Sequence 512 AA:

Query Match 58.6%; Score 51; DB 22; Length 512;

Best Local Similarity 62.5%; Pred. No. 12;

Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Or 1 MSLTVFLLIKWTG 16

Db 1 mslatvllwgtly 16

RESULT 3

AAB66894 standard; Protein: 23 AA.

AC AAB66894:

XX 12-APR-2001 (first entry)

XX Human M019 signal sequence.

XX Human: M010; M019; M024; cancer; obesity; diabetes; osteoporosis; asthma;

XX nervous system disorder; pain; stroke; hypercholesterolemia; atherosclerosis;

XX chronic obstructive pulmonary disease; hyperlipidemia; arteriosclerosis;

XX hyperlipoproteinemia; arteriosclerosis; coronary artery disease.

XX Homo sapiens.

XX MO200100644-A1.

XX 04-JAN-2001.

XX 30-JUN-2000; 2000MO-US18097.

XX 30-JUN-1999; 9905-0345293.

XX 30-JUN-1999; 9905-0345680.

XX (MILL-) MILLENNium PHARM INC.

XX Khodadoust M;

XX WPI: 2001-112434/12.

XX N-PSDB: AAF32084, AAF32085.

XX New gene encoding proteins designated M010, M019, M024, useful for the

XX diagnosis, prophylaxis and treatment of cancer, obesity, diabetes,

XX osteoporosis, asthma, atherosclerosis and pulmonary disorders -

XX Claim 8: Page 93: 98pp: English.

XX The present invention relation relates to novel human coding sequences
 CC and proteins designated M010 (AAF32082 and AAB66884), M019 (AAF32084 and
 CC AAB66893) and M024 (AAF32086 and AAB66896). M010, M019 and M024 proteins
 CC and compounds which modulate their activity or which bind M010, M019
 CC and/or M024 are useful for prevention and treatment of a variety of
 CC central and peripheral nervous system disorders, pain, stroke, chronic
 CC obstructive pulmonary disease, hypercholesterolemia, hyperlipidemia,
 CC hyperlipoproteinemia, arteriosclerosis, arteriosclerosis and coronary
 CC artery disease. The present sequence is the human M019 protein signal
 CC sequence.

SO Sequence 23 AA:

Query Match 55.2%; Score 48; DB 22; Length 23;

Best Local Similarity 72.7%; Pred. No. 1.2;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Or 6 VILLIKWTGK 16

Db 12 villikwtgk 22

RESULT 4

AAB66893 standard; Protein: 85 AA.

XX AAB66893:

XX 12-APR-2001 (first entry)

XX Human M019 protein.

XX Human: M010; M019; M024; cancer; obesity; diabetes; osteoporosis; asthma;

XX nervous system disorder; pain; stroke; hypercholesterolemia; atherosclerosis;

XX chronic obstructive pulmonary disease; hyperlipidemia; arteriosclerosis;

XX hyperlipoproteinemia; arteriosclerosis; coronary artery disease.

XX Homo sapiens.

XX MO200100644-A1.

XX 04-JAN-2001.

XX 30-JUN-2000; 2000MO-US18097.

XX 30-JUN-1999; 9905-0345293.

XX 30-JUN-1999; 9905-0345680.

XX (MILL-) MILLENNium PHARM INC.

XX Khodadoust M;

XX WPI: 2001-112434/12.

XX N-PSDB: AAF32084, AAF32085.

XX New gene encoding proteins designated M010, M019, M024, useful for the

XX diagnosis, prophylaxis and treatment of cancer, obesity, diabetes,

XX osteoporosis, asthma, atherosclerosis and pulmonary disorders -

XX Disclosure: Fig 2: 98pp: English.

XX The present invention relation relates to novel human coding sequences

XX and proteins designated M010 (AAF32082 and AAB66884), M019 (AAF32084 and

XX AAB66893) and M024 (AAF32086 and AAB66896). M010, M019 and M024 proteins

PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134222.
PR 14-MAY-1999; 9905-0134223.
PR 14-MAY-1999; 9905-0134268.
PR 19-MAY-1999; 9905-0134641.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 26-MAY-1999; 9905-0136192.
PR 27-MAY-1999; 9905-0136222.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 11-JUN-1999; 9905-0138583.
PR 14-JUN-1999; 9905-0138119.
PR 16-JUN-1999; 9905-0139453.
PR 16-JUN-1999; 9905-0139453.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139493.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139463.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140354.
PR 29-JUN-1999; 9905-0140961.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142390.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144321.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144632.
PR 21-JUL-1999; 9905-0144632.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145086.
PR 22-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145193.
PR 23-JUL-1999; 9905-0145145.

PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145216.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145913.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146386.
PR 03-AUG-1999; 9905-0146386.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147200.
PR 04-AUG-1999; 9905-0147200.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147933.
PR 10-AUG-1999; 9905-0148170.
PR 11-AUG-1999; 9905-0148170.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149366.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149425.
PR 18-AUG-1999; 9905-0149425.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150888.
PR 26-AUG-1999; 9905-0150888.
PR 27-AUG-1999; 9905-0151063.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152663.
PR 10-SEP-1999; 9905-0152746.
PR 10-SEP-1999; 9905-0152746.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155135.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155929.
PR 24-SEP-1999; 9905-0155929.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158025.
PR 08-OCT-1999; 9905-0158322.
PR 12-OCT-1999; 9905-0158563.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159294.
PR 14-OCT-1999; 9905-0159320.
PR 14-OCT-1999; 9905-0159320.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 14-OCT-1999; 9905-0159638.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.


```
PR 22-OCT-1999: 9905-0160981-
PR 22-OCT-1999: 9905-0160989-
PR 22-OCT-1999: 9905-0161404-
PR 22-OCT-1999: 9905-0161405-
PR 22-OCT-1999: 9905-0161406-
PR 26-OCT-1999: 9905-0161359-
PR 26-OCT-1999: 9905-0161360-
PR 26-OCT-1999: 9905-0161361-
PR 26-OCT-1999: 9905-0161920-
PR 28-OCT-1999: 9905-0161992-
PR 28-OCT-1999: 9905-0161993-
PR 29-OCT-1999: 9905-0161942-

Query Match 50.6% Score 44: DB 21: Length 200:
Best Local Similarity 60.0% Pred. No. 45:
Matches 9: Conservative 1: Mismatches 5: Gaps 0:

OY 2 SMLVWLLNGVYTKG 16
    | | | | | | | | |
DB 171 stlslslgslgslg 185

RESULT 9
AACG28189
ID AAG28189 standard: Protein: 228 AA.
AC AACG28189:
CX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33315.
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termintation sequence.
KM Arabidopsis thaliana.
XX EPI03405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000: 2000EP-0301439.
XX 25-FEB-1999: 9905-0121825.
XX 05-MAR-1999: 9905-0123180.
XX 09-MAR-1999: 9905-0123548.
XX 23-MAR-1999: 9905-0125788.
XX 25-MAR-1999: 9905-0126264.
XX 29-MAR-1999: 9905-0126785.
XX 01-APR-1999: 9905-0127452.
XX 08-APR-1999: 9905-0128714.
XX 16-APR-1999: 9905-0128845.
XX 19-APR-1999: 9905-0130077.
XX 21-APR-1999: 9905-0130449.
XX 23-APR-1999: 9905-0130510.
XX 23-APR-1999: 9905-0130891.
XX 28-APR-1999: 9905-0131449.
XX 30-APR-1999: 9905-0132407.
XX 04-MAY-1999: 9905-0132484.
XX 05-MAY-1999: 9905-0132485.
XX 06-MAY-1999: 9905-0132486.
XX 07-MAY-1999: 9905-0132487.
XX 11-MAY-1999: 9905-0132853.
XX 14-MAY-1999: 9905-0134218.
XX 14-MAY-1999: 9905-0134221.
XX 14-MAY-1999: 9905-0134370.
XX 18-MAY-1999: 9905-0134768.
XX 19-MAY-1999: 9905-0134941.

PR 20-MAY-1999: 9905-0135124.
PR 21-MAY-1999: 9905-0135353.
PR 24-MAY-1999: 9905-0135659.
PR 25-MAY-1999: 9905-0136071.
PR 26-MAY-1999: 9905-0136072.
PR 01-JUN-1999: 9905-0136282.
PR 03-JUN-1999: 9905-0137222.
PR 04-JUN-1999: 9905-0137502.
PR 07-JUN-1999: 9905-0137724.
PR 08-JUN-1999: 9905-0138094.
PR 10-JUN-1999: 9905-0138540.
PR 14-JUN-1999: 9905-0139119.
PR 16-JUN-1999: 9905-0139452.
PR 17-JUN-1999: 9905-0139492.
PR 18-JUN-1999: 9905-0139454.
PR 18-JUN-1999: 9905-0139455.
PR 18-JUN-1999: 9905-0139456.
PR 18-JUN-1999: 9905-0139457.
PR 18-JUN-1999: 9905-0139458.
PR 18-JUN-1999: 9905-0139459.
PR 18-JUN-1999: 9905-0139460.
PR 18-JUN-1999: 9905-0139461.
PR 18-JUN-1999: 9905-0139462.
PR 18-JUN-1999: 9905-0139463.
PR 18-JUN-1999: 9905-0139750.
PR 22-JUN-1999: 9905-0139852.
PR 22-JUN-1999: 9905-0139859.
PR 23-JUN-1999: 9905-0140353.
PR 23-JUN-1999: 9905-0140354.
PR 24-JUN-1999: 9905-0140695.
PR 26-JUN-1999: 9905-0140823.
PR 29-JUN-1999: 9905-0140993.
PR 30-JUN-1999: 9905-0141384.
PR 01-JUL-1999: 9905-0141842.
PR 02-JUL-1999: 9905-0142154.
PR 06-JUL-1999: 9905-0142390.
PR 08-JUL-1999: 9905-0142803.
PR 09-JUL-1999: 9905-0142920.
PR 12-JUL-1999: 9905-0143512.
PR 14-JUL-1999: 9905-0143624.
PR 15-JUL-1999: 9905-0144085.
PR 16-JUL-1999: 9905-0144086.
PR 16-JUL-1999: 9905-0144325.
PR 19-JUL-1999: 9905-0144331.
PR 19-JUL-1999: 9905-0144333.
PR 19-JUL-1999: 9905-0144334.
PR 19-JUL-1999: 9905-0144335.
PR 20-JUL-1999: 9905-0144353.
PR 20-JUL-1999: 9905-0144631.
PR 20-JUL-1999: 9905-0144884.
PR 21-JUL-1999: 9905-0144814.
PR 21-JUL-1999: 9905-0145088.
PR 22-JUL-1999: 9905-0145088.
PR 22-JUL-1999: 9905-0145089.
PR 22-JUL-1999: 9905-0145089.
PR 22-JUL-1999: 9905-0145197.
PR 23-JUL-1999: 9905-0145145.
PR 23-JUL-1999: 9905-0145218.
PR 26-JUL-1999: 9905-0145276.
PR 27-JUL-1999: 9905-0145913.
PR 27-JUL-1999: 9905-0145918.
PR 28-JUL-1999: 9905-0145919.
PR 28-JUL-1999: 9905-0145951.
```

PR 02-AUG-1999; 9905-0146386.
 PR 02-AUG-1999; 9905-0146388.
 PR 02-AUG-1999; 9905-0146389.
 PR 03-AUG-1999; 9905-0147038.
 PR 03-AUG-1999; 9905-0147039.
 PR 04-AUG-1999; 9905-0147302.
 PR 05-AUG-1999; 9905-0147192.
 PR 05-AUG-1999; 9905-0147260.
 PR 06-AUG-1999; 9905-0147303.
 PR 06-AUG-1999; 9905-0147416.
 PR 06-AUG-1999; 9905-0147493.
 PR 07-AUG-1999; 9905-0148171.
 PR 10-AUG-1999; 9905-0148173.
 PR 11-AUG-1999; 9905-0148319.
 PR 12-AUG-1999; 9905-0148341.
 PR 13-AUG-1999; 9905-0148565.
 PR 13-AUG-1999; 9905-0148684.
 PR 16-AUG-1999; 9905-0149358.
 PR 17-AUG-1999; 9905-0149415.
 PR 18-AUG-1999; 9905-0149420.
 PR 20-AUG-1999; 9905-0149723.
 PR 20-AUG-1999; 9905-0149929.
 PR 23-AUG-1999; 9905-0149902.
 PR 23-AUG-1999; 9905-0149930.
 PR 25-AUG-1999; 9905-0150566.
 PR 26-AUG-1999; 9905-0150884.
 PR 27-AUG-1999; 9905-0151060.
 PR 27-AUG-1999; 9905-0151066.
 PR 30-AUG-1999; 9905-0151303.
 PR 31-AUG-1999; 9905-0151438.
 PR 01-SEP-1999; 9905-0151930.
 PR 07-SEP-1999; 9905-0152363.
 PR 07-SEP-1999; 9905-0152368.
 PR 13-SEP-1999; 9905-0153758.
 PR 15-SEP-1999; 9905-0154018.
 PR 16-SEP-1999; 9905-0154039.
 PR 20-SEP-1999; 9905-0154779.
 PR 22-SEP-1999; 9905-0155139.
 PR 23-SEP-1999; 9905-0155486.
 PR 24-SEP-1999; 9905-0156559.
 PR 24-SEP-1999; 9905-0156560.
 PR 26-SEP-1999; 9905-0156596.
 PR 04-OCT-1999; 9905-0157117.
 PR 05-OCT-1999; 9905-0157753.
 PR 06-OCT-1999; 9905-0157865.
 PR 07-OCT-1999; 9905-0158029.
 PR 08-OCT-1999; 9905-0158232.
 PR 12-OCT-1999; 9905-0158369.
 PR 12-OCT-1999; 9905-0158370.
 PR 13-OCT-1999; 9905-0158294.
 PR 13-OCT-1999; 9905-0159295.
 PR 14-OCT-1999; 9905-0159329.
 PR 14-OCT-1999; 9905-0159330.
 PR 14-OCT-1999; 9905-0159331.
 PR 14-OCT-1999; 9905-0159637.
 PR 14-OCT-1999; 9905-0159638.
 PR 14-OCT-1999; 9905-0159639.
 PR 21-OCT-1999; 9905-0160741.
 PR 21-OCT-1999; 9905-0160767.
 PR 21-OCT-1999; 9905-0160768.
 PR 21-OCT-1999; 9905-0160770.
 PR 21-OCT-1999; 9905-0160814.
 PR 21-OCT-1999; 9905-0160815.
 PR 22-OCT-1999; 9905-0160880.
 PR 22-OCT-1999; 9905-0160881.
 PR 22-OCT-1999; 9905-0160989.
 PR 25-OCT-1999; 9905-0161404.
 PR 25-OCT-1999; 9905-0161405.
 PR 25-OCT-1999; 9905-0161359.
 PR 26-OCT-1999; 9905-0161360.

PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161992.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.
 PR 29-OCT-1999; 9905-0162142.
 Query Match 50.6%; Score 44; DB 21; Length 228;
 Best Local Similarity 60.0%; Pred. No. 52;
 Matches 9; Conservative 1; Mismatches 5; Gaps 0;
 DB 199 stslslsl9w9tllg 213
 OY 2 SMWVFLMAGCTWAG 16
 I I I I I I I I I I
 RESULT 10
 ID AAY13399
 AC AAY13399 standard; Protein: 280 AA.
 XX
 XX AAY13399;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of protein PRO268.
 XX
 KW Secreted protein; transmembrane protein; human; enterocolitis;
 KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
 KW congenital ulcerovillous atrophy; skin disease; cell growth;
 KW skin disease; skin disease; skin disease; skin disease;
 KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
 KW fibromodulin; dermal scarring; Usher Syndrome; Atrophila areata;
 KW anti-thrombotic; wound healing; tissue repair.
 XX
 OS Homo sapiens.
 XX
 XX W09914328-42.
 XX
 PD 25-MAR-1999.
 XX
 PF 16-SEP-1998; 98WO-US19330.
 XX
 XX 25-NOV-1997; 97US-0066840.
 PR 17-SEP-1997; 97US-0059113.
 PR 17-SEP-1997; 97US-0059114.
 PR 17-SEP-1997; 97US-0059117.
 PR 17-SEP-1997; 97US-0059119.
 PR 17-SEP-1997; 97US-0059121.
 PR 17-SEP-1997; 97US-0059122.
 PR 17-SEP-1997; 97US-0059184.
 PR 18-SEP-1997; 97US-0059263.
 PR 18-SEP-1997; 97US-0059265.
 PR 18-SEP-1997; 97US-0059266.
 PR 17-OCT-1997; 97US-0062285.
 PR 17-OCT-1997; 97US-0062287.
 PR 21-OCT-1997; 97US-0063486.
 PR 24-OCT-1997; 97US-0062814.
 PR 24-OCT-1997; 97US-0062816.
 PR 24-OCT-1997; 97US-0063045.
 PR 24-OCT-1997; 97US-0063121.
 PR 24-OCT-1997; 97US-0063127.
 PR 24-OCT-1997; 97US-0063128.
 PR 27-OCT-1997; 97US-0063327.
 PR 27-OCT-1997; 97US-0063421.
 PR 28-OCT-1997; 97US-0063544.
 PR 28-OCT-1997; 97US-0063547.
 PR 28-OCT-1997; 97US-0063549.
 PR 28-OCT-1997; 97US-0063550.
 PR 28-OCT-1997; 97US-0063564.
 PR 29-OCT-1997; 97US-0063435.
 PR 29-OCT-1997; 97US-0063704.
 PR 29-OCT-1997; 97US-0063732.

XX (INCY-) INCYTE PHARM INC.
 XX Lai P, Guegler KJ, Gorgone GA, Corley KC, Baughn MR, Tang YH,
 XX Williams JC, Bandman O, Azimzal Y, Au-Young J, Tse H, Lu DM;
 XX Yang J
 DR WPI: 2000-303785/26.
 DR N-PSDB; AAA03375.
 XX Purified polypeptide for treating or preventing disorders associated
 XX with decreased expression or activity of oxidoreductase molecules
 PS Claim 1; Page 69-70; 97pp; English.
 XX
 CC AA92504-18 show OXRE-1 to -15; OXRE-1, an oxidoreductase, has identity
 CC with chloroexin.
 CC The polypeptides are useful for treating or preventing a disorder
 CC associated with decreased expression or activity of OXRE. Antagonists
 CC of OXRE are useful for treating or preventing a disorder associated with
 CC decreased expression or activity of OXRE.
 CC proliferative disorders (cirrhosis, hepatitis) cancer (leukemia, cell
 CC melanoma), hypopituitarism and hyperpituitarism, hypochromidism and
 CC hyperchromidism, metabolic disorders (Addison's disease, cystic
 CC fibrosis), reproductive disorders (infertility, ovulatory defects),
 CC neurological disorders (Alzheimer's disease, Parkinson's disease,
 CC multiple sclerosis), mental disorders (anxiety, schizophrenia),
 CC immunological disorders (acquired immunodeficiency syndrome
 CC (AIDS), asthma, osteoarthritis) and other infectious diseases.
 CC polynucleotides may be used in Southern or Northern analysis, polymerase
 CC chain reaction (PCR), or in enzyme-linked immunosorbent assays (ELISA).
 XX
 XX Sequence 280 AA:
 S0
 Query Match 50.6%; Score 44; DB 21; Length 280;
 Best Local Similarity 53.8%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 07 3 MLVPELLLCVTV 15
 DB 11 lavlvllllvgapw 23
 RESULT 14
 AAA40021
 ID AAA40021 standard; Protein: 280 AA.
 XX
 AC AAA40021;
 XX
 XX 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 3166.
 XX
 XX Human; noctlopic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amphotropic; lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX neurotrophic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukemia.
 OS Homo sapiens.
 XX
 XX MO20015312-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 26-DEC-2000; 2000MO-US34263.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0692036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSC-) HYSCD INC.
 XX
 XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao GA, Zhou P, Goodrich R, Drmanac RT;
 DR WPI: 2001-442253/47.
 DR N-PSDB; AA159177.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 XX
 XX Example 4; SEQ ID NO 3166; 10078pp; English.
 XX
 XX The invention relates to human nucleic acids (AA157796-AA161369) and
 XX polypeptides (AA161370-AA161371). The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX centralised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amphotropic
 XX lateral sclerosis, and Shy-Drager syndrome. Other uses include the
 XX treatment of immunological disorders, such as acquired immunodeficiency
 XX syndrome (AIDS), asthma, osteoarthritis and other infectious diseases.
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukemias and
 XX C.N.S. disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 280 AA:
 S0
 Query Match 50.6%; Score 44; DB 22; Length 280;
 Best Local Similarity 53.8%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 07 3 MLVPELLLCVTV 15
 DB 11 lavlvllllvgapw 23
 RESULT 15
 AAB88357
 ID AAB88357 standard; Protein: 280 AA.
 XX
 AC AAB88357;
 XX
 XX 23-MAY-2001 (first entry)
 DE Human membrane or secretory protein clone P58C0083.
 XX
 XX Human; secretory protein; membrane protein; vaccine; gene therapy;
 XX tricuspid aortic architis; diabetes.
 XX
 OS Homo sapiens.
 XX
 XX EP1067182-A2.
 XX
 XX 10-JAN-2001.
 XX
 XX 07-JUL-2000; 2000EP-0114090.
 XX
 XX 08-JUL-1999; 99JP-0194179.
 XX 11-JAN-2000; 2000JP-0118775.
 XX 02-MAY-2000; 2000JP-0183766.

PA (HELI-) HELIX RES INST.

PL Ota T, Isogai T, Mishikawa T, Kawai Y, Sugiyama T, Hayashi K;

PR WPI: 2001-093869/11.

DR N-TSD01: AAF93784.

PT Nucleic acids encoding secretory proteins/membrane proteins, useful in gene therapy or as candidate target molecules in drug development -

XX Claim 1: SEQ ID 82; 609pp + CD ROM: English.

XX This invention relates to nucleic acid sequences AAF93744 - AAF93916 which encode human secretory or membrane proteins represented by

CC AAF89317 - AAF88419. Included in the invention are primers

CC AAF94295 and AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the invention. The invention also includes methods for the production of antibodies directed against the proteins, and cDNA sequences, which can be used in vaccines. The polynucleotide sequences can be used in gene therapy. The polynucleotide sequences and the proteins encoded by the nucleic acid sequences may be used in the diagnosis of diseases associated with inappropriate secretory and protein/membrane protein expression. The nucleic acids and complementary sequences may also be used as DNA probes to detect and quantitate the presence of similar nucleic acid sequences in samples. They may also be used to study the expression and function of secretory proteins/membrane proteins. The polynucleotide sequences and the polypeptides may be used as antigens in the production of antibodies. The polypeptides may be used as identity modulators (agonists and antagonists) of expression and activity. The antibodies and antagonists may also be used as therapeutic agents to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the polypeptides in samples (e.g. by enzyme linked immunosorbent assay (ELISA) Experiments of diseases which may be treated include rheumatoid arthritis and diabetes.

CC Sequence 280 AA:

50

Query Match 50.6%; Score 44; DB 22; Length 280;

Best Local Similarity 53.8%; Pctn. No. 64;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

0Y 3 MLVFFLLMGVTW 15

DB 11 Iavilvllwgsppw 23

RESULT 16

AAAB80267

ID AAB80267 standard: Protein: 280 AA.

XX AAB80267:

XX 24-APR-2001 (first entry)

XX Human PRO266 protein.

XX Human: PRO: dermatological; antiproliferic; cytostatic; antiinflammatory;

KM antiparkinsonian neurologic; neuroprotective; vulnerary; cancer;

KM antiangiogenic; vasotrophic; antihemorrhagic; antineurologic; cancer;

KM antischistosomal; antileishmanial; antidiabetic; antiviral; diabetes;

KM optineurologic; gene therapy; skin disease; gastrointestinal disorder;

XX Ischemia; inflammation.

OS Homo sapiens.

XX

XX W0200104311-A1.

XX 18-JAN-2001.

XX 22-FEB-2000, 2000MO-US04414.

XX 07-JUL-1999; 9905-0143048.

XX 26-JUL-1999; 9905-0145698.

XX 28-JUL-1999; 9905-0146222.

XX 16-SEP-1999; 99MO-US20394.

XX 16-SEP-1999; 99MO-US20394.

XX 15-SEP-1999; 99MO-US21090.

XX 15-SEP-1999; 99MO-US21547.

XX 05-OCT-1999; 99MO-US23089.

XX 29-NOV-1999; 99MO-US28214.

XX 30-NOV-1999; 99MO-US28313.

XX 16-DEC-1999; 99MO-US30093.

XX 20-DEC-1999; 99MO-US30393.

XX 05-JAN-2000; 99MO-US00219.

PA (GENENTECH INC.

XX Ashkenazi AJ, Botstein D, Desnovers L, Eaton DL, Ferrara N,

P1 Pflavret E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A,

P1 Methan JP, Pao L, Pao L, Pao L, Pao L, Pao L, Pao L, Pao L,

P1 Williams PM, Wood WT.

XX WPI: 2001-081051/09.

DR N-PSDB: AAF72429.

XX Sixty one nucleic acids encoding PRO polypeptides which are useful in gene therapy or as candidate target molecules in drug development -

PT squamous cell carcinoma) and neurodegenerative diseases (e.g. Alzheimer's disease) -

XX Claim 1: Fig 114; 393pp: English.

XX The present sequence is one of sixty one novel secreted and useful for the treatment of the diseases (e.g. Alzheimer's disease, Parkinson's disease, wound repair, cardiovascular disorders (e.g. endometrial bleeding, angioneurosis, ischemias such as coronary ischemia, atherosclerosis), inflammatory disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis), interstitial, AIDS and diabetes mellitus).

CC The PRO nucleic acids have applications in molecular biology, including use as hybridization probes, and in chromosome and gene mapping.

XX Sequence 280 AA:

50

Query Match 50.6%; Score 44; DB 22; Length 280;

Best Local Similarity 53.8%; Pctn. No. 64;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

0Y 3 MLVFFLLMGVTW 15

DB 11 Iavilvllwgsppw 23

RESULT 17

AAV25756

ID AAV25756 standard: Protein: 281 AA.

XX AAV25756:

XX 04-OCT-1999 (first entry)

XX Human secreted protein encoded from gene 46.

XX

XX Secreted protein; human; treatment; diagnosis; therapy; cancer; tumor;

KM neurodegenerative disorder; developmental abnormality; blood disorder;

KM fetal deficiency; blood disorder; leukemia; immune system; inflammation;

KM autoimmune disease; hepatic disease; renal disease; allergy; restenosis;

KM ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
 KM cardiovascular disorder; wound healing; stroke; arthritis; obesity;
 KM asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
 KM metabolic disorder.
 OS Homo sapiens.
 XX W09398881-AL.
 XX
 XX 05-AUG-1999.
 XX
 XX 27-JAN-1999. 99MO-US01621.
 XX
 XX 30-JAN-1998. 98US-0073170.
 PR 30-JAN-1998. 98US-0073159.
 PR 30-JAN-1998. 98US-0073160.
 PR 30-JAN-1998. 98US-0073161.
 PR 30-JAN-1998. 98US-0073162.
 PR 30-JAN-1998. 98US-0073164.
 PR 30-JAN-1998. 98US-0073165.
 PR 30-JAN-1998. 98US-0073167.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Carter KC, Endress GA, Feng P, Ferrie AM, Florence G;
 PI Florence KA, Janet F, Ni J, Rosen CA, Ruben SM;
 PI Soppet DR, Young P, Yu G.
 PI MPI: 1999-46915/39.
 N-PSDB: AM200455.
 XX
 XX New isolated human genes and the secreted polypeptides they encode
 PT useful in, e.g. treatment of Alzheimer's
 XX
 XX Claim 1b: Page 323-324; 393pp: English.
 XX
 XX This invention describes novel human genes (see AA200410-200477) and the
 CC secreted proteins (see AA25711-Y25778) and fragments (see
 CC AA25779-Y25907) they encode. The polynucleotides and their corresponding
 CC secreted polypeptides are useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. Also pathological
 CC conditions can be diagnosed by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 67
 CC genes disclosed in this invention, including the use of the genes for
 CC highly expressed in, and include developing products for the diagnosis
 CC or treatment of cancer, tumours, neurodegenerative disorders, disorders
 CC of developmental abnormalities and fetal deficiencies, blood disorders,
 CC leukemias, diseases of the immune system, autoimmune diseases, hepatic
 CC and renal diseases, inflammation, allergies, ischaemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, arthritis, obesity, asthma, sepsis,
 CC disorders, wound healing, stroke, arthritis, asthma, cardiovascular
 CC disease, AIDS, metabolic disorder, infection, transplant rejection, ischaemic
 CC and AIDS. The polypeptides are also useful for identifying their binding
 XX partners.
 XX
 XX Sequence 281 AA:
 SO
 Query Match 50.6%, Score 44; DB 20; Length 281;
 Best Local Similarity 53.8%, Pval N=62;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0.

XX
 XX 17-OCT-2000 (first entry)
 DT
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 33314.
 DE
 XX protein identification: signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX
 XX Arabidopsis thaliana.
 OS
 XX EP103405-A2.
 XX
 XX 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999. 99US-0121835.
 PR 05-MAR-1999. 99US-0123160.
 PR 05-MAR-1999. 99US-0123168.
 PR 05-MAR-1999. 99US-0123169.
 PR 25-MAR-1999. 99US-0126264.
 PR 25-MAR-1999. 99US-0126785.
 PR 01-APR-1999. 99US-0127462.
 PR 06-APR-1999. 99US-0128234.
 PR 08-APR-1999. 99US-0128714.
 PR 16-APR-1999. 99US-0129845.
 PR 19-APR-1999. 99US-0130077.
 PR 23-APR-1999. 99US-0130101.
 PR 23-APR-1999. 99US-0130691.
 PR 28-APR-1999. 99US-0131449.
 PR 30-APR-1999. 99US-0132048.
 PR 04-MAY-1999. 99US-0132407.
 PR 04-MAY-1999. 99US-0132484.
 PR 05-MAY-1999. 99US-0132485.
 PR 06-MAY-1999. 99US-0132487.
 PR 07-MAY-1999. 99US-0132663.
 PR 11-MAY-1999. 99US-0134256.
 PR 14-MAY-1999. 99US-0134218.
 PR 14-MAY-1999. 99US-0134219.
 PR 14-MAY-1999. 99US-0134221.
 PR 14-MAY-1999. 99US-0134222.
 PR 14-MAY-1999. 99US-0134223.
 PR 19-MAY-1999. 99US-0134941.
 PR 20-MAY-1999. 99US-0135124.
 PR 21-MAY-1999. 99US-0135353.
 PR 24-MAY-1999. 99US-0135629.
 PR 25-MAY-1999. 99US-0136021.
 PR 27-MAY-1999. 99US-0136392.
 PR 28-MAY-1999. 99US-0136782.
 PR 03-JUN-1999. 99US-0137526.
 PR 04-JUN-1999. 99US-0137502.
 PR 07-JUN-1999. 99US-0137724.
 PR 08-JUN-1999. 99US-0138094.
 PR 10-JUN-1999. 99US-0138540.
 PR 10-JUN-1999. 99US-0138847.
 PR 14-JUN-1999. 99US-0139113.
 PR 15-JUN-1999. 99US-0139115.
 PR 16-JUN-1999. 99US-0139453.
 PR 17-JUN-1999. 99US-0139459.
 PR 18-JUN-1999. 99US-0139454.
 PR 18-JUN-1999. 99US-0139455.
 PR 18-JUN-1999. 99US-0139456.
 PR 18-JUN-1999. 99US-0139457.
 PR 18-JUN-1999. 99US-0139458.
 PR 18-JUN-1999. 99US-0139459.
 PR 18-JUN-1999. 99US-0139460.
 PR 18-JUN-1999. 99US-0139461.
 PR 18-JUN-1999. 99US-0139462.
 PR 18-JUN-1999. 99US-0139463.

RESULT 18
 AACG28188
 ID AACG28188 standard. Protein: 312 AA.
 XX
 AC AACG28188:

DB 3 MLVFFLLKNGVTM 15
 II IAVLVIIIGSPW 23

PR 18-JUN-1999 9905-0139750.
 PR 18-JUN-1999 9905-0139763.
 PR 18-JUN-1999 9905-0139784.
 PR 22-JUN-1999 9905-0139899.
 PR 23-JUN-1999 9905-0140353.
 PR 23-JUN-1999 9905-0140354.
 PR 24-JUN-1999 9905-0140695.
 PR 28-JUN-1999 9905-0140823.
 PR 29-JUN-1999 9905-0140993.
 PR 01-JUL-1999 9905-0141189.
 PR 01-JUL-1999 9905-0142154.
 PR 02-JUL-1999 9905-0142055.
 PR 06-JUL-1999 9905-0142390.
 PR 08-JUL-1999 9905-0142803.
 PR 09-JUL-1999 9905-0142920.
 PR 12-JUL-1999 9905-0142977.
 PR 13-JUL-1999 9905-0143624.
 PR 14-JUL-1999 9905-0143627.
 PR 15-JUL-1999 9905-0144008.
 PR 16-JUL-1999 9905-0144085.
 PR 16-JUL-1999 9905-0144086.
 PR 19-JUL-1999 9905-0144325.
 PR 19-JUL-1999 9905-0144331.
 PR 19-JUL-1999 9905-0144332.
 PR 19-JUL-1999 9905-0144333.
 PR 19-JUL-1999 9905-0144334.
 PR 19-JUL-1999 9905-0144335.
 PR 20-JUL-1999 9905-0144352.
 PR 20-JUL-1999 9905-0144632.
 PR 20-JUL-1999 9905-0144884.
 PR 21-JUL-1999 9905-0144883.
 PR 21-JUL-1999 9905-0145088.
 PR 22-JUL-1999 9905-0145085.
 PR 22-JUL-1999 9905-0145087.
 PR 22-JUL-1999 9905-0145089.
 PR 22-JUL-1999 9905-0145192.
 PR 23-JUL-1999 9905-0145145.
 PR 23-JUL-1999 9905-0145146.
 PR 23-JUL-1999 9905-0145276.
 PR 26-JUL-1999 9905-0145276.
 PR 27-JUL-1999 9905-0145913.
 PR 27-JUL-1999 9905-0145918.
 PR 27-JUL-1999 9905-0145919.
 PR 28-JUL-1999 9905-0145951.
 PR 02-AUG-1999 9905-0146386.
 PR 02-AUG-1999 9905-0146388.
 PR 02-AUG-1999 9905-0146389.
 PR 03-AUG-1999 9905-0147038.
 PR 04-AUG-1999 9905-0147204.
 PR 04-AUG-1999 9905-0147302.
 PR 05-AUG-1999 9905-0147192.
 PR 05-AUG-1999 9905-0147260.
 PR 06-AUG-1999 9905-0147302.
 PR 06-AUG-1999 9905-0147303.
 PR 09-AUG-1999 9905-0147493.
 PR 09-AUG-1999 9905-0147935.
 PR 10-AUG-1999 9905-0148171.
 PR 11-AUG-1999 9905-0148319.
 PR 12-AUG-1999 9905-0148341.
 PR 13-AUG-1999 9905-0148565.
 PR 13-AUG-1999 9905-0148684.
 PR 14-AUG-1999 9905-0148688.
 PR 17-AUG-1999 9905-0149175.
 PR 18-AUG-1999 9905-0149426.
 PR 20-AUG-1999 9905-0149722.
 PR 20-AUG-1999 9905-0149723.
 PR 23-AUG-1999 9905-0149929.
 PR 23-AUG-1999 9905-0149902.
 PR 23-AUG-1999 9905-0150290.
 PR 26-AUG-1999 9905-0150884.
 PR 26-AUG-1999 9905-0150884.

PR 27-AUG-1999 9905-0151055.
 PR 27-AUG-1999 9905-0151065.
 PR 27-AUG-1999 9905-0151080.
 PR 27-AUG-1999 9905-0151081.
 PR 31-AUG-1999 9905-0151438.
 PR 01-SEP-1999 9905-0151930.
 PR 07-SEP-1999 9905-0152363.
 PR 10-SEP-1999 9905-0153070.
 PR 13-SEP-1999 9905-0153756.
 PR 15-SEP-1999 9905-0154018.
 PR 15-SEP-1999 9905-0154019.
 PR 20-SEP-1999 9905-0154778.
 PR 22-SEP-1999 9905-0155139.
 PR 23-SEP-1999 9905-0155486.
 PR 24-SEP-1999 9905-0155659.
 PR 28-SEP-1999 9905-0156456.
 PR 29-SEP-1999 9905-0156599.
 PR 04-OCT-1999 9905-0157114.
 PR 05-OCT-1999 9905-0157115.
 PR 06-OCT-1999 9905-0157862.
 PR 07-OCT-1999 9905-0158028.
 PR 08-OCT-1999 9905-0158232.
 PR 12-OCT-1999 9905-0158369.
 PR 13-OCT-1999 9905-0159293.
 PR 13-OCT-1999 9905-0159294.
 PR 13-OCT-1999 9905-0159295.
 PR 14-OCT-1999 9905-0159330.
 PR 14-OCT-1999 9905-0159331.
 PR 14-OCT-1999 9905-0159637.
 PR 14-OCT-1999 9905-0159638.
 PR 18-OCT-1999 9905-0159584.
 PR 21-OCT-1999 9905-0160741.
 PR 21-OCT-1999 9905-0160742.
 PR 21-OCT-1999 9905-0160768.
 PR 21-OCT-1999 9905-0160770.
 PR 21-OCT-1999 9905-0160811.
 PR 21-OCT-1999 9905-0160815.
 PR 22-OCT-1999 9905-0160980.
 PR 22-OCT-1999 9905-0160981.
 PR 22-OCT-1999 9905-0160983.
 PR 25-OCT-1999 9905-0161404.
 PR 25-OCT-1999 9905-0161405.
 PR 25-OCT-1999 9905-0161406.
 PR 26-OCT-1999 9905-0161359.
 PR 26-OCT-1999 9905-0161360.
 PR 26-OCT-1999 9905-0161361.
 PR 28-OCT-1999 9905-0161929.
 PR 28-OCT-1999 9905-0161930.
 PR 28-OCT-1999 9905-0161931.
 PR 29-OCT-1999 9905-0162142.

Query Match 50.6% Score 44: DB 21: Length 312:
 Best Local Similarity 60.0% Pred. No. 72:
 Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:
 Oy 2 SMLVFFLLKMGKNG 16
 Db 283 sflasfljgwjclg 297

RESULT 19
 AAB23623
 ID AAB23623 standard: Protein: 372 AA.
 AC AAB23623:
 XY
 XX 12-JAN-2001 (first entry)
 XX
 DE Human secreted protein SKO ID NO: 46.
 XX
 XX Human, secreted protein; cytokine; cell proliferation;
 XX nutritional supplement; immune modulation; autoimmune disorder.
 XX

KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 XX Peptide 16..28
 XX PT /label= signal_peptide
 XX FT 29..372
 XX Protein /label= mature_protein
 PF MO200049134-A1.
 XX 24-AUG-2000.
 XX 18-FEB-2000: 2000MO-US04340.
 XX 19-FEB-1999: 990US-0120680.
 XX 23-APR-1999: 990US-0298733.
 XX 17-AUG-1999: 990US-0149639.
 XX 03-SEP-1999: 990US-0155685.
 XX 25-NOV-1999: 990US-0167823.
 XX 29-NOV-1999: 990US-0167823.
 XX 15-FEB-2000: 2000US-0298733.
 XX (ALPH-) ALPHAGEN INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P:
 XX WPI: 2000-549267/50.
 XX N-PSDB: AAB3123.
 XX New secreted proteins and polynucleotides encoding them, which are
 XX derived from Homo sapiens, useful for therapy, diagnosis, and research,
 XX as well as nutritional sources or supplements.
 XX Claim 55: Page 272-274, 309pp; English.
 XX The present sequence is the sequence of a human secreted protein. Its
 XX cDNA was isolated from an adult lung cDNA library. The proteins
 XX and coding sequences of the invention can be used in the isolation of
 XX similar genes and proteins. In the elucidation of their function *in vivo*,
 XX and to treat a number of conditions. It is possible that they may have
 XX immunomodulatory activity. The proteins of the invention may be used to
 XX factors, in immune modulation, where they may be used to treat immune and
 XX autoimmune diseases, as haematopoiesis regulators (treating myeloid or
 XX lymphoid cell deficiencies). In the promotion of tissue growth, they may
 XX have chemokine or chemotactic activity, haemostatic or thrombolytic
 XX activity, or anti-inflammatory activity.
 XX Sequence 372 AA:
 SQ
 Query Match 50.6%; Score 44; DB 21; Length 372;
 Best Local Similarity 53.3%; Pred. No. 87;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 SMYVFIILKGYTNG 16
 DB 164 svltaflvlpwmg 178
 RESULT 20
 AAB23643
 ID AAB23643 standard; Protein: 421 AA.
 XX AAB23643:
 XX 12-JAN-2001 (first entry)
 XX Human secreted protein SFO ID NO: 99.
 XX Human secreted protein; cytokine; cell proliferation;

KW nutritional supplement; immune modulation; autoimmune disorder;
 KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 XX Peptide 16..28
 XX PT /label= signal_peptide
 XX FT 29..372
 XX Protein /label= mature_protein
 PF MO200049134-A1.
 XX 24-AUG-2000.
 XX 18-FEB-2000: 2000MO-US04340.
 XX 19-FEB-1999: 990US-0120680.
 XX 23-APR-1999: 990US-0298733.
 XX 17-AUG-1999: 990US-0149639.
 XX 03-SEP-1999: 990US-0155686.
 XX 29-NOV-1999: 990US-0157247.
 XX 29-NOV-1999: 990US-0167823.
 XX 15-FEB-2000: 2000US-0298733.
 XX (ALPH-) ALPHAGEN INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P:
 XX WPI: 2000-549267/50.
 XX New secreted proteins and polynucleotides encoding them, which are
 XX derived from Homo sapiens, useful for therapy, diagnosis, and research,
 XX as well as nutritional sources or supplements.
 XX Disclosure: Page 304-305, 309pp; English.
 XX The present invention is concerned with a number of secreted proteins
 XX and their coding sequences isolated from various human cDNA libraries.
 XX The proteins and coding sequences can be used in the isolation of
 XX similar genes and proteins. In the elucidation of their function *in vivo*,
 XX and to treat a number of conditions. It is possible that they may have
 XX immunomodulatory activity. The proteins of the invention may be used to
 XX factors, in immune modulation, where they may be used to treat immune and
 XX autoimmune diseases, as haematopoiesis regulators (treating myeloid or
 XX lymphoid cell deficiencies). In the promotion of tissue growth, they may
 XX have chemokine or chemotactic activity, haemostatic or thrombolytic
 XX activity, or anti-inflammatory activity. No information about sequences
 XX AAB23643-B23645 is given in the specification.
 XX Sequence 421 AA:
 SQ
 Query Match 50.6%; Score 44; DB 21; Length 421;
 Best Local Similarity 53.3%; Pred. No. 99;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 SMYVFIILKGYTNG 16
 DB 164 svltaflvlpwmg 178
 RESULT 21
 AAY57899
 ID AAY57899 standard; Protein: 439 AA.
 XX AAY57899:
 XX 23-MAR-2000 (first entry)
 XX Human transmembrane protein HTPMPN-23.
 XX Human: transmembrane protein; HTPMPN; diagnosis; immunospecific;
 XX antiproliferative; neuroprotective; immune disorder;
 XX reproductive disorder; smooth muscle disorder; neurological disorder;
 XX gastrointestinal disorder; developmental disorder;
 KW cell proliferative disorder.

OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
PH	Misc-difference 436
FT	/note: "unspecified"
XX	
PN	M09661471-AZ.
PD	
PD	02-DEC-1999.
PE	
PE	28-MAY-1999; 99MO-US11904.
PR	
PR	20-MAY-1998; 98US-008726D.
PR	02-JUL-1998; 98US-009167A.
PR	02-OCT-1998; 98US-0102954.
PR	24-NOV-1998; 98US-0109869.
PA	(INCYTE -) INCYTE PHARM INC.
PI	
PI	Tong WT, Lai P, Hillman JT, Yue H,
PI	Bendman O, Patterson C, Gorgone CA,
PI	Au-Yang J, Kaser MK, Braughn MK;
PT	
PT	WPI: 2000-072605/06.
DR	M-PISDB: AA256720.
DR	
PS	Proteins, polynucleotides, vectors, host cells and antibodies used to
PS	prepare vaccines prevent immune, reproductive, smooth muscle
PS	neurological, gastrointestional, developmental and coll proliferative
PS	disorders -
PS	Claim 1: Page 130-131; 229pg: English.
XX	
XX	AA256698 to AA25776 encode AAVS7877 to AAVS7955 which represent human
XX	transmembrane proteins designated hTRM-1 to hTRM-79, respectively.
XX	The amino acid sequence of each protein is set forth in the following and
XX	neuropeptide activities. The human transmembrane proteins
CC	polynucleotide encoding them and other compositions and methods from
CC	the present invention can be used for the diagnosis, treatment or
CC	prevention of immune, reproductive, smooth muscle, neurological,
CC	gastrointestinal, developmental and coll proliferative disorders. The
CC	hTRM-1 can be used to treat or prevent disorders associated with a
CC	decreased expression or activity of hTRM-1.
XX	
XX	Sequence 439 AA:
SO	
Query Match	50.6%; Score 44; DB 21; Length 439;
Best Local Similarity	53.3%; Pred. No. 1e-03;
Matches	8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Oy	2 SMLYRLLIMQYMG 16
	I:I I:I:I:I I
Db	164 avtlatvltwphmg 178
RESULT: 22	
AAU04131	
AC	AAU04131 standard: Protein; 541 AA.
XX	
XX	AAU04131:
XX	
DJ	24-OCT-2001 (first entry)
DE	
DE	Human novel protein #2.
XX	
XX	Human novel protein. Antinuclear, osteopathic, antiinflammatory;
KM	immunomodulatory; osteoblast, osteogenic; antiproliferative; C-
KM	anticonvulsant; antihistatic; cerebroprotective; antifungal; antiviral;
KM	antibacterial; antiallergic; dermatoprotective; haemostatic; antiasumatic;
KM	chromolytic; immunogen; antibody; gene therapy; neurological disorder;
KM	Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KM	tissue regeneration; immune disorder.

XX	Homo sapiens.
OS	M02001554.37 -A2.
PX	02-ANG-2001.
XX	25-JAN-2001; 2001MO-US02653.
PF	25-JAN-2000; 2000US-0491404.
PR	(HUSE) HUSE INC.
PI	Tang YT, Liu C, Dmanac RT:
PT	WPI: 2001-451939/48.
DR	N-PSTB: AAS22436.
XX	Isolated polypeptides useful for treating anti-inflammatory diseases,
PT	nervous system disorders, and for regenerating bone and cartilage -
XX	Example 4; Page 520-521; 694pp; English.
CC	The invention relates to polynucleotides encoding novel human
CC	proteins or their active domains. The polypeptides, polynucleotides and
CC	antibodies raised against the polypeptides are used in a method of
CC	treatment of a mammal and prevention of disorders caused by the aberrant
CC	protein expression or activity. The polypeptides can be used as
CC	immunogens to elicit antibodies which bind to the protein for detection.
CC	The polypeptides are used to identify compounds which bind to the
CC	polypeptides. Polynucleotides of the invention are used as probes and
CC	primers, for sequencing, for chromosome or gene mapping. In the
CC	production of recombinant proteins, and in generating anti-sense DNA or
CC	RNA and in gene therapy. Polypeptides of the invention can be used to
CC	target drugs to a tumour. In assays to determine biological activity, to
CC	test antitumoural agents, to determine immune response, to determine quantitation
CC	of antigen-antibody complexes, to study immunological processes, to study
CC	disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC	platelets of the invention may also be useful in treating platelet
CC	disorders, ligament and/or nerve tissue, wound healing, treating burns, promoting
CC	the proliferation, differentiation and survival of stem cells, as a
CC	contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC	Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC	sclerosis, infection of from autoimmunity causing allergy, asthma, viral or
CC	fungal infection of host disease, eczema, haemophilia, thrombosis, cancer,
CC	anti-inflammatory diseases, nervous system disorders, and infection.
CC	The present sequence represents a protein of the invention.
XX	Sequence 541 AA:
QY	Query Match 50.6%; Score 44; DN 22; Length 541;
DB	Best Local Similarity 53.3%; Prd. NO. 1.3e+02;
Matches	8; Conservative 2; Mismatches 5; Indels 0; Gaps 0
OY	2 SMLVVELLALNGYVG 16 164 SVLEATLVLDPhmg 178
RESULT 23	
AAC89162	
1D	AAG89162 standard; Protein: 541 AA.
XX	AAG89162:
XX	11-SEP-2001 (first entry)
DE	Human secreted protein, SEQ ID NO: 282.
XX	Human: secreted protein; gene vaccine; treatment; diagnosis;
XX	GENSET.

XX OS Homo sapiens.
 XX PN W0200142451-A2.
 XX 14-JUN-2001.
 XX PF 07-DEC-2000; 2000MO-1B01938.
 XX 08-DEC-1999; 9905-0169629.
 XX 06-MAR-2000; 200005-0187470.
 XX (GENSET) GENSET.
 XX Dumas Mline Edwards J, Bougueleret L, Jobert S;
 XX WPI: 2001-367870/38.
 XX N-PSDB: AAH64765.
 XX Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases -
 XX Claim 21: Page 817-818; 921pp: English.
 XX The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides of diseases associated with inappropriate GENSET gene expression. For example, they are used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patient's own production of GENSET polypeptides. Conversely, antisenase nucleic acid molecules may be administered to down regulate GENSET expression. The sense and antisenase nucleic acids may also be used to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antidiodes and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention.
 XX Sequence 541 AA:
 XX
 XX Query Match 50.6%; Score 44; DB 22; Length 541;
 XX Best Local Similarity 53.3%; Pred. No. 1,3e+02;
 XX Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 XX
 XX 2 SMUVEFLMGCTNG 16
 XX 164 svlatfivlpwpg 178
 XX
 XX RESULT 24
 XX AAG89180 standard: Protein: 541 AA.
 XX AAG89180:
 XX 11-SEP-2001 (first entry)
 XX Human secreted protein, SED ID NO: 300.
 XX Human: secreted protein: gene therapy: vaccine: treatment: diagnosis:
 XX GENSET.
 XX Homo sapiens.
 XX W0200142451-A2.
 XX

PD 14-JUN-2001.
 XX PF 07-DEC-2000; 2000MO-1B01938.
 XX 08-DEC-1999; 9905-0169629.
 XX 06-MAR-2000; 200005-0187470.
 XX (GENSET) GENSET.
 XX Dumas Mline Edwards J, Bougueleret L, Jobert S;
 XX WPI: 2001-367870/38.
 XX N-PSDB: AAH64783.
 XX Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases -
 XX Claim 21: Page 829-831; 921pp: English.
 XX The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides of diseases associated with inappropriate GENSET gene expression. For example, they are used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patient's own production of GENSET polypeptides. Conversely, antisenase nucleic acid molecules may be administered to down regulate GENSET expression. The sense and antisenase nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy. The GENSET polypeptides may be used as antigens in the production of antidiodes and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention.
 XX Sequence 541 AA:
 XX
 XX Query Match 50.6%; Score 44; DB 22; Length 541;
 XX Best Local Similarity 53.3%; Pred. No. 1,3e+02;
 XX Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 XX
 XX 2 SMUVEFLMGCTNG 16
 XX 164 svlatfivlpwpg 178
 XX
 XX RESULT 25
 XX AAY01453 standard: Protein: 40 AA.
 XX AAY01453:
 XX 18-MAY-1999 (first entry)
 XX Secreted protein encoded by gene 63 clone HMAN70.
 XX Human: secreted protein: gene therapy: protein: cancer: weight:
 XX tumour: chromosome mapping: forensic: hematological disease: allergy:
 XX inflammation: cell proliferation: viral infection: wound healing:
 XX modulation: appetite: behaviour: food additive: preservative.
 XX Homo sapiens.
 XX W09903990-A1.
 XX 28-JAN-1999.
 XX 15-JUL-1998; 98MO-US14613.
 XX

KM GL protein; diagnostic; vaccine; equine viral arteritis.
 OS Equine arteritis virus.
 XX MO9519438-A1.
 PN 20-JUL-1995.
 XX 13-JAN-1995: 95WO-GH00066.
 PF 14-JAN-1994: 94GB-0000656.
 XX (UKAC-) UK MIN FISHERIES & FOOD.
 PA Chlorniside ED:
 XX WPI: 1995-263868/34.
 DR N-PSDB; AA098814.
 XX New equine arteritis virus peptide(s), conjugates, DNA and
 PF antibodies - are used as vaccines or diagnostic agents in equine
 PT arteritis virus-mediated diseases
 PS Claim 1: Page 18; 31pp; English.
 XX The DNA sequence equivalent to the entire EAV genome minus the
 CC first 18 bases and the poly-A tail is given in AA098814, and
 CC includes the GL protein coding sequence. Epitopes of GL are
 CC used in vaccine protin.
 XX Sequence 137 AA:
 50
 Query Match 49.4%; Score 43; DB 16; Length 137;
 Best Local Similarity 58.3%; Pred. No. 42;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 CY 1 MEMVYVFLIING 12
 DB 2 LAMIVLIFLIING 13
 RESULT 28
 AA74573 standard; Protein: 218 AA.
 XX AA74573:
 AC AA74573:
 XX 21-MAR-2000 (first entry)
 DT
 XX Neisseria gonorrhoeae ORF 152 protein sequence SPO ID NO:620.
 DE
 XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 OS Neisseria gonorrhoeae.
 XX MO9557280-A2.
 PN 11-NOV-1999.
 XX 30-APR-1999: 95WO-US09346.
 PF 01-MAY-1998: 98US-0083758.
 PR 31-JUL-1998: 98US-0094869.
 PR 02-SEP-1998: 98US-0098994.
 PR 02-SEP-1998: 98US-0098994.
 PR 02-SEP-1998: 98US-0098994.
 PR 09-OCT-1998: 98US-0103749.
 PR 09-OCT-1998: 98US-0103749.
 PR 09-OCT-1998: 98US-0103749.
 PR 25-FEB-1999: 99US-0121528.
 XX

PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
 PI Petersen J, Pizzo M, Rappelli R, Ratti G, Scatato E, Scarselli M,
 XX Tettelin H, Venter JC;
 DR WPI: 2000-062150/05.
 DR N-PSDB; AA53335.
 XX Novel Neisserial polypeptides predicted to be useful antigens for
 PF vaccines and diagnostics
 PT
 XX Claim 2: Page 435; 1453pp; English.
 XX AA53015 to AA254536, AA254577 to AA254615, and AA74253 to AA75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunologic compositions. The polypeptides can be used in the
 CC manufacture of polypeptide vaccines, in the prevention of infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC may also be used in gene therapy protocols.
 XX Sequence 218 AA:
 50
 Query Match 49.4%; Score 43; DB 21; Length 218;
 Best Local Similarity 43.8%; Pred. No. 68;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 CY 1 MEMVYVFLIING 16
 DB 43 VGLIVLIFLIVFLIING 58
 RESULT 29
 AA74574 standard; Protein: 218 AA.
 XX AA74574:
 XX 21-MAR-2000 (first entry)
 DT
 XX Neisseria meningitidis ORF 152 protein sequence SPO ID NO:622.
 DE
 XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 OS Neisseria meningitidis.
 XX MO9557280-A2.
 PN 11-NOV-1999.
 XX 30-APR-1999: 95WO-US09346.
 PF 01-MAY-1998: 98US-0083758.
 PR 31-JUL-1998: 98US-0094869.
 PR 02-SEP-1998: 98US-0098994.
 PR 02-SEP-1998: 98US-0098994.
 PR 02-SEP-1998: 98US-0098994.
 PR 09-OCT-1998: 98US-0103749.
 PR 09-OCT-1998: 98US-0103749.
 PR 09-OCT-1998: 98US-0103749.
 PR 25-FEB-1999: 99US-0121528.
 XX (CHIR) CHIRON CORP.

Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition -

polymorphic variants AAF33095 - AAP3142 encode human secreted proteins AAB64773 - AAB64820. Fragments of the secreted proteins and amino acid sequences which share homology with the fragments are represented in AAB64821 - AAB64880. The genes and proteins have been deposited to their tissues and cells in which they are expressed. The secreted proteins are involved in the regulation of hair growth and antagonists include: immunosuppressive, antiarthritic, antirheumatic; antiproliferative; cytostatic; cardiant; vasoprotic; cerebroprotective; neurotropic; neuroprotective; antibacterial; vitiligo; fungicide; and ophthalmological. The secreted proteins, polymorphoids, antagonists and agonists may be useful in treating, preventing and diagnosing diseases and disorders such as autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis, Crohn's disease, ulcerative colitis, breast or liver cardiovascular disorders e.g. cardiac arrest, atherosclerosis, cerebrovascular disorders e.g. cerebral ischaemia, and degenerative nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria viruses and fungi and ocular disorders e.g. corneal infection. The polyphosphates can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn to maintain organs before transplantation, for supporting cell culture to primary tissues, used as a food additive or preservative, as a microcapsule storage capabilities. Included in the invention are sequences AAB64772 and AAF33095 - AAP3142 which are used in the isolation and characterisation of the nucleotide and protein sequences of the invention.

Sequence 37 AA;

Query Match	48.38;	Score 42;	DB 22;	Length 37;
Best Local Similarity	63.68;	Pred. No. 15;		
Matches	7; Conservative	2; Mismatches	2; Indels	0; Gaps

y	2	SMLVFLLWG	12
		1:1	11:111
b	4	SLIAFLLWG	14

RESULT 32

D AAB76763 standard; Protein; 402 AA

AAB76763;

11-APR-2001 (first entry)

Corynebacterium glutamicum MCT protein SEQ ID NO:508

genome mapping; genetic engineering.

Corynebacterium glutamicum.

MO200100805-A2

04-JAN-2001

23-JUN-2000; 2000WG-1B00926

R	25-JUN-1999;	99US-0141031
R	08-JUL-1999;	99DE-1031454
R	08-JUL-1999;	99DE-1031478
R	08-JUL-1999;	99DE-1031563
R	09-JUL-1999;	99DE-1032122

[illegible]

(BADI) BASF AG

PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G
XX
DR MPI; 2001-071486/08.

DR N-PSDB; AAF67996

PT Cozymetacellum glutamicum nucleic acids encoding membrane construction proteins or their portions, useful for typing and membrane transport proteins or their portions, useful for typing on PT identifying C. glutamicum or related bacteria, and as markers for PT transformation -

PS Claim 20; Page 868-870; 1119pp; English.

CC AAE6770 to AAE6800 encode the *Corynebacterium glutamicum* membrane
CC construction and membrane transport (MCT) proteins given in AAB75510 to
CC AAE6770 to AAE6800. The amino acid sequence of the MCT proteins was
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in *C. glutamicum* or
CC related bacteria (e.g. *Brevibacterium lactofermentum*), the typing or
CC identification of *C. glutamicum* or related bacteria, as reference points
CC AAE6802 and AAE6802 represent sequencing primers which are used in an
CC example from the present invention.

SQ Sequence 402 AA.

Query Match	48.38;	Score 42;	DB 22;	length 402;
Best Local Similarity	50.08;	Pred. No. 1.8e+02;		
Matches	6;	Conservative	3;	Mismatches 3; Indels 0; Gaps 0

```
QY      4  LVVFLLLMGVTW 15
          :|:| | | :|
Db      377  lviIglawgIkw 38
```

RESULT 33

ID AAG92801 standard; Protein; 449 AA

AC AAG92801;

XX DT 26-SEP-2001 (first entry)
 XX DB C glutamicum protein fragment SEQ ID NO: 6555.
 XX CC Corynebacterium bacterium; amino acid synthesis; vitamin; saccharide;
 XX KW organic acid synthetase.
 XX OS Corynebacterium glutamicum.
 XX BP108790-AZ.
 XX 20-JUN-2001.
 XX PF 18-DEC-2000: 2000BP-0127688.
 XX PR 16-DEC-1999: 99JP-0377484.
 XX PR 07-APR-2000: 2000JP-0159162.
 XX PR 03-AUG-2000: 2000JP-0280988.
 XX RA (KYO) KYOKA HAKKO KOGYO KK.
 XX RA NAGASAWA S, MIZOGUCHI H, ANDO S, HAYASHI M, OCHIAI K, YOKOI H;
 XX PI TATEISHI N, Senoh A, Ikeda M, Ozaki A;
 XX WP: 2001-376931/40.
 XX N-PSDB: AMH68020.
 XX DB Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT -
 XX Claim 17: SEQ ID NO: 6555; 246pp + Sequence Listing: English.
 XX CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from corynebacterium bacterium, and identifying a homologue of a gene derived
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, organic acids, vitamins, and other compounds.
 CC particularly L-valine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX CC
 XX Sequence 449 AA:
 XX
 XX Query Match 48.3%; Score 42; DB 22; Length 449;
 XX Best Local Similarity 50.0%; Pred. No. 2e+02;
 XX Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX 07 4 LUYELLKQYTW 15
 XX ::::|:::|:::|
 XX DB 424 ivtfglwgkfw 435
 XX
 XX RESULT 34
 XX AAK12362
 XX ID AAK12362 standard; Protein: 455 AA.
 XX AAK12362;
 XX 13-SBP-1991 (first entry)
 XX OCtopus rhodopsin membrane protein.
 XX DE OCtopus; rhodopsin; membrane; helix; OR.
 XX KW
 XX XX

PN JP03123486-A.
 XX XX
 XX PD 27-MAY-1991.
 XX XX
 XX PF 06-OCT-1989; 89JP-0260261.
 XX XX
 XX PF 06-OCT-1989; 89JP-0260261.
 XX XX
 XX PA (HITA) HITACHI KK.
 XX XX
 XX WP: 1991-197925/27.
 XX N-PSDB: AAO12225.
 XX PT Genetic engineering of membrane protein - by division of protein
 PT into cartridge genes corresp. to helix structure polypeptides)
 XX XX
 XX PS Disclosure: Fig 1; 17pp; Japanese.
 XX XX
 XX CC A gene cassette is prep. by division of the OR membrane protein
 CC gene encoding helix structure polypeptides. A base sequence contg.
 CC the coding region of the gene is amplified by PCR using primers
 CC from the 6' coll. tryptophan operon. The region of the tripe polypeptide
 CC transition regulating base sequence, and the N-terminal Met of the tripe
 CC (or tripe) or tripe polypeptide, respectively, is introduced into
 CC an expression vector for transformation of host cells. A partial
 CC or total OR membrane protein is produced by culturing the
 CC transformants.
 XX CC
 XX Sequence 455 AA:
 XX
 XX Query Match 48.3%; Score 42; DB 12; Length 455;
 XX Best Local Similarity 23.1%; Pred. No. 2.1e+02;
 XX Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX 07 3 MUYELLKQYTW 15
 XX ::::|:::|:::|
 XX DB 155 imlvfwmsisw 167
 XX
 XX RESULT 35
 XX AAY09518
 XX ID AAY09518 standard; Protein: 596 AA.
 XX AAY09518;
 XX 16-JUL-1999 (first entry)
 XX CC C. elegans dopamine transport protein CeDAT1.
 XX DE C. elegans; dopamine transport protein; CeDAT1; CeDAT1; nematode;
 XX KW antiparasitic; psychoactive drug.
 XX OS Caenorhabditis elegans.
 XX XX
 XX W09921883-A1.
 XX 06-MAY-1999.
 XX PF 27-OCT-1998; 98WO-0522712.
 XX XX
 XX PR 27-OCT-1997; 97US-0063282.
 XX XX
 XX PA (AMCY) AMERICAN CYANAMID CO.
 XX (UTVA-) UNIV WANDERBILT.
 XX Blakely RD, Eppeler CM;
 XX WP: 1999-312943/26.
 XX N-PSDB: AAX5244.
 XX DB Novel C.elegans dopamine transporter proteins
 XX PT
 XX XX

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases and cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

	CC	measure the binding or interaction of a ligand (e.g. gabapentin,
	CC	L-Norleucine, L-Allo-Isoleucine, L-methionine, L-leucine,
	CC	L-Valine, Spermine and/or L-phenylalanine) of a calcium channel
	CC	alpha2delta subunit.
	XX	
SQ	Sequence	304 AA:
	Query Match	47.7% Score 41.5; DB 22: Length 304;
	Best Local Similarity	69.2%; Pred. No. 1.6e+02;
	Matches	9; Conservative 1; Mismatches 2; Indels 1; Gaps 1
OY	4 IYVFTLLKGYTNG 16	
Db	209 IYIITLWVSF-Wg 220	
	RESULT 41	
AAB6Z244		
ID	AAB6Z244 standard: Protein: 304 AA.	
AA		
AA	AAB6Z244:	
PT	11-JUN-2001 (first entry)	
DE	Human calcium channel alpha2delta subunit related seq ID No. 16.	
NM	Calcium channel alpha2delta subunit: alpha2delta-1; cerebral cortex;	
NM	nervous system disorder: pain, epilepsy, anxiety, human.	
OS	Homo sapiens.	
PX	M0200120336-A2.	
PD	22-MAR-2001.	
PR	18-SEP-2000: 2000ONO-EPO9136.	
PR	16-SEP-1999: 99OUS-0397549.	
PA	(WARN) WARNER LAMBERT CO.	
PI	Bertelli F, Brown JP, Dissanayake V, Suman-Chauhan N, Gee NS;	
XX	WPT: 2001-357902/26.	
DR	M-PDB: AAT57547.	
PT	Competitive binding assay for screening ligands which bind a cerebral	
PT	where the ligands identified are useful for treating disorders of the	
PT	nervous system, including pain .	
XX		
XX	Disclosure: Page 84: 15pp: English.	
XX	The invention relates to a new method for screening ligands which bind a	
CC	cerebral cortical voltage-dependent calcium channel alpha2delta subunit,	
CC	preferably alpha2delta-1 subunit. The method comprises contacting a	
CC	secreted soluble recombinant alpha2delta-1 subunit with a ligand of	
CC	interest and a labelled compound which binds the subunit, followed by	
CC	measuring the level of binding of the labelled compound to alpha2delta-1	
CC	biologically active products that modulate a nervous system function,	
CC	which bind a cerebral cortical voltage-dependent calcium channel	
CC	alpha2delta-1 subunit. The ligands identified by the method are useful	
CC	for treating disorders of the nervous system, including pain, epilepsy	
CC	and anxiety. The present sequence represents a human calcium channel	
XX	alpha2delta subunit related sequence.	
XX		
XX	Sequence 304 AA:	
Query Match	47.7% Score 41.5; DB 22: Length 304;	
Best Local Similarity	69.2%; Pred.No. 1.6e+02;	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2002, 16:49:55 (Search time 19.31 seconds)

18.646 Million cell updates/sec (without alignments)

Title: US-09-471-276-831_COPY_1_16

Perfect score: 87

Sequence: 1 KSMUYPTLLMGVYWG 16

Scoring table: BLCSUM62

Gapop 10.0, Dapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database: Issued Patents.AI.*

1: /cgn2_6/ptodata/2/19a/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/19a/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/19a/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/19a/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/19a/6C_COMB.pep.*

6: /cgn2_6/ptodata/2/19a/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Length	DB	ID	Description
1	44	50.6	280	1	US-09-564-119C-2	Sequence 2, App11
2	43	49.4	137	1	US-09-576-169-2	Sequence 2, App11
3	40	46.0	503	4	US-09-068-195-24	Sequence 24, App11
4	39	44.8	277	1	US-08-024-868-2	Sequence 2, App11
5	39	44.8	277	2	US-08-442-097-2	Sequence 2, App11
6	39	44.8	277	4	US-08-500-095-2	Sequence 1, App11
7	39	44.8	339	3	US-09-120-365-80	Sequence 62, App11
8	39	44.8	339	3	US-09-120-365-80	Sequence 62, App11
9	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
10	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
11	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
12	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
13	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
14	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
15	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
16	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
17	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
18	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
19	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
20	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
21	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
22	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
23	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
24	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
25	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
26	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11
27	39	44.8	339	4	US-09-515-039-62	Sequence 62, App11

ALIGNMENTS

```

RESULT      1
US-09-264-419C-2
: Sequence 2: Application US/09264419C
: Patent No. 6174682
: GENERAL INFORMATION:
: APPLICANT: Rhododouel, Mehtrun
: TITLE OF INVENTION: No. 6174682el Thiorodoxin Family Active Site Molecules and Uses
: FILE REFERENCE: MNT-0765
: CURRENT FILING DATE: 1999-03-08
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO: 2
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-264-419C-2

Query Match      50.6%; Score 44: DB 4: Length 280:
Best Local Similarity 53.8%; Freq. No. 28:
Matches 17: Conservative 2: Mismatches 4: Indels 0: Gaps 0:

OY      3  MLVPELLMGVTW 15
DB      11  LAVILLMGAPW 23

RESULT      2
US-08-676-169-2
: Sequence 2: Application US/08676169
: Patent No. 5773235
: GENERAL INFORMATION:
: APPLICANT: Chirnside, Ewan Douglas
: TITLE OF INVENTION: EQUINE ARTERIAL VIRUS PEPTIDES, ANTIBODIES
: FILE REFERENCE: MNT-0765
: CURRENT FILING DATE: 1997-03-13
: NUMBER OF SEQ ID NOS: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHIVE P.C.
: STREET: 1100 NO. 5773235th Gable Road, 8th floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22201-5714
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/676.169
: FILING DATE: 1997-03-13
: CLASSIFICATION: A35
: ATTORNEY/AGENT INFORMATION:
: NAME: Crawford, Arthur R.
: REGISTRATION NUMBER: 25,327
: REFERENCE/DOCKET NUMBER: 1498-85
: TELEPHONE: (703) 816-4006
: TELEFAX: (703) 816-4100
: INFORMATION FOR SEQ ID NO: 2:
: LENGTH: 137 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-676-169-2

```

```

Query Match      49.4%; Score 43: DB 1: Length 137:
Best Local Similarity 58.3%; Freq. No. 19:
Matches 17: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

OY      1  MSNVLVFLMG 12
DB      2  LSNVLFLMG 13

RESULT      3
US-09-068-195-24
: Sequence 2: Application US/09068195B
: Patent No. 6140078
: GENERAL INFORMATION:
: APPLICANT: Sanders, Jan W.
: APPLICANT: Ledebner, Adriaans M.
: APPLICANT: Venema, Gerard
: TITLE OF INVENTION: Salt-Inducible Premoter Derivable from a Lactate Acid
: TITLE OF INVENTION: Bacterially and Immunologically Active Lactate Acid Bacterium for
: FILE REFERENCE: Sanders-6013/0282227
: CURRENT APPLICATION NUMBER: US/09/068.195B
: EARLIER FILING DATE: 1998-07-29
: EARLIER APPLICATION NUMBER: PCY/EP97/04795
: EARLIER FILING DATE: 1997-03-13
: EARLIER APPLICATION NUMBER: EP 9620244/7
: EARLIER FILING DATE: 1997-03-13
: EARLIER APPLICATION NUMBER: EP 9620244/4
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO: 44
: TYPE: PRT
: ORGANISM: L. lactis MG1363
US-09-068-195-24

Query Match      46.0%; Score 40: DB 4: Length 503:
Best Local Similarity 46.7%; Freq. No. 18+02:
Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:

OY      4  LVFPLMLMGVTW 15
DB      35  LVFPLMLGGLW 46

RESULT      4
US-08-024-868-2
: Sequence 2: Application US/08024868
: Patent No. 5386013
: GENERAL INFORMATION:
: APPLICANT: Lee, Tae Ho
: APPLICANT: Wisniewski, Hans Georg
: APPLICANT: Vliet, Jan
: TITLE OF INVENTION: Tyrosine-Induced Protein, TSC-6, DNA
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brovady and Nelmark
: STREET: 419 seventh street, NW
: CITY: Washington
: STATE: DC
: COUNTRY: US
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.24

```



```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/024,868
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
PRIORITY DATE: 13-MAY-1994
APPLICATION NUMBER: US/07/642,312
FILING DATE: 13-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Litvack, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: VILCEK-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-024-868-2

```

```

Query Match 44.8%; Score 39; DB 1; Length 277;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;
OY 3 MIVPFLIMQV--WG 16
DB 4 LIYFLIMEDTQGMG 19

```

```

RESULT 5
US-08-242-097-2
Sequence 2, Application US/08242097
Patent No. 5846729
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Misiewicz, Hans Georg
TITLE OF INVENTION: Cytokine-Induced Protein, TSC-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
ADDRESS: Broadway and Melmark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/242,097
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Litvack, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-242-097-2

```

```

Query Match 44.8%; Score 39; DB 2; Length 277;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;
OY 3 MIVPFLIMQV--WG 16
DB 4 LIYFLIMEDTQGMG 19

```

```

RESULT 6
US-09-206-695-2
Sequence 2, Application US/09206695
Patent No. 6210905
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Misiewicz, Hans Georg
TITLE OF INVENTION: Cytokine-Induced Protein, TSC-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
ADDRESS: Broadway and Melmark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/206,695
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Litvack, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-206-695-2

```

```

Query Match 44.8%; Score 39; DB 4; Length 277;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;
OY 3 MIVPFLIMQV--WG 16
DB 4 LIYFLIMEDTQGMG 19

```

```

RESULT 7
PCT-US96-11995-1
Sequence 1, Application PCT/7059611995
Patent No. 5846729
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Misiewicz, Hans Georg
TITLE OF INVENTION: Cytokine-Induced Protein, TSC-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
ADDRESS: Broadway and Melmark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/206,695
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Litvack, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
TELEPHONE: 212-737-3528
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-206-695-2

```

;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Broadway and Melmark
;; STREET: 419 Seventh Street, N.W., Suite 300
;; CITY: Washington
;; STATE: D.C.
;; ZIP: 20004
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT FILING DATE: 1997-11-18
;; APPLICATION NUMBER: PCT/US96/11995
;;
;; FILING DATE:
;; CLASSIFICATION:
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 60/001,311
;; FILING DATE: 20-JUL-1995
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SHUNJI
;; REGISTRATION NUMBER: 25,618
;; REFERENCE/DOCKET NUMBER: MISNEMSKI-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-628-5197
;; TELEFAX: 212-737-3528
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 277 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
;; MOLECULE TYPE: protein
;; PRT-US96-11995-1

Query Match 44.8%; Score 39; DB 5; Length 277;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 3 MLVEELLGCVT--WG 16
DB 4 LTVLPFLIMEDTGWG 19
:::|||||111

RESULT 8
US-09-120-365-62
; Sequence 62, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 80
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
US-09-120-365-62

Query Match 44.8%; Score 39; DB 3; Length 339;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 KQWTWG 16
DB 308 KQWTWG 313
:::|||||

RESULT 9
US-09-120-365-80
; Sequence 80, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120,365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 80
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
US-09-120-365-80

Query Match 44.8%; Score 39; DB 3; Length 339;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 KQWTWG 16
DB 308 KQWTWG 313
:::|||||

RESULT 10
US-09-515-039-62
; Sequence 62, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 62
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Sarcophaga cathepsin L
US-09-515-039-62

Query Match 44.8%; Score 39; DB 4; Length 339;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 KQWTWG 16
DB 308 KQWTWG 313
:::|||||

RESULT 11
US-09-515-039-80
; Sequence 80, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 80
 LENGTH: 339
 TYPE: PRT
 ORGANISM: Sarcophaga caltropis L
 US-09-515-035-80

Query Match 44.8%; Score 39; DB 4; Length 339;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 MOVING 16
 DB 308 MOVING 313

RESULT 12

US-08-681-192-2
 ; Sequence 2, Application US/08681192
 ; Patent No. 6281811
 ; GENERAL INFORMATION:
 ; APPLICANT: BERGSM, DEBK
 ; APPLICANT: SARAU, HENRY
 ; APPLICANT: RUBEN, STEVEN
 ; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR HNF0578
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESSES:
 ; CORRESPONDENCE ADDRESSEE: Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-2199
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/681.192
 ; FILING DATE: 23-NOV-1993
 ; PRIORITY APPLICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Schreck, Patricia A.
 ; REGISTRATION NUMBER: 33,777
 ; REFERENCE/DOCKET NUMBER: AT050014
 ; TELEPHONE: 610-270-4021
 ; TELEFAX: 610-270-4026
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 344 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; US-08-681-192-2

Query Match 44.8%; Score 39; DB 4; Length 344;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 SMLVPELW 11
 DB 242 AIVWPELW 251

RESULT 13

US-08-445-640-4
 ; Sequence 4, Application US/08445640
 ; Patent No. 5709858
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Marx, Melanie R.
 ; APPLICANT: Scadden, David T.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Barton, WILLI F.
 ; TITLE OF INVENTION: Protein Tyrosine Kinases
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/445.640
 ; FILING DATE: 22-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/170558
 ; PRIORITY APPLICATION: 1993
 ; APPLICATION DATA: 08/157563
 ; FILING DATE: 23-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hasak, Janet E.
 ; REGISTRATION NUMBER: 28,616
 ; REFERENCE/DOCKET NUMBER: B54C2
 ; TELEPHONE: 415/225-1986
 ; TELEFAX: 415/252-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 913 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-445-640-4

Query Match 44.8%; Score 39; DB 1; Length 913;
 Best Local Similarity 38.5%; Pred. No. 4.4e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 MIVWPELW 15
 DB 431 LIIWPELW 443

RESULT 14

US-08-170-558-4
 ; Sequence 4, Application US/08170558
 ; Patent No. 6001621
 ; GENERAL INFORMATION:
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Marx, Melanie R.
 ; APPLICANT: Scadden, David T.
 ; APPLICANT: Baker, Kevin P.

```

? APPLICANT: Barton, Will F.
? TITLE OF INVENTION: Protein Tyrosine Kinases
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS: 35
? INVENTOR: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? PRIORITY DATE: 23-NOV-1993
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Palin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/170,558
? FILING DATE: 20-DEC-1993
? CLASSIFICATION: 435
? PRIORITY NUMBER: 433
? APPLICATION NUMBER: 08/157563
? FILING DATE: 23-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Haasak, Janet E.
? REGISTRATION NUMBER: 28,616
? REFERENCE/DOCKET NUMBER: 854C1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415/225-1896
? TELEFAX: 415/952-9881
? TELETYPE: 910/371-7168
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 913 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
?
? US-08-170-558-4
?
? Query Match 44.8%; Score 39; DB 3; Length 913;
? Best Local Similarity 38.5%; Pred. No. 4,4e+02;
? Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
?
? QY 3 MVLVPLILNGYTW 15
? : : : : : : : : : : : : : : : : : :
? DB 431 LLLIALMLRLHML 443
?
? RESULT 15
? US-08-447-314-4
? Sequence 4, Application US/08447314
? GENERAL INFORMATION:
? PATENT NO. 6096527
? APPLICANT: Scadden, David T.
? APPLICANT: Baker, Kevin P.
? APPLICANT: Barton, Will F.
? TITLE OF INVENTION: Protein Tyrosine Kinases
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESS: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? PRIORITY DATE: 23-NOV-1993
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Palin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/447,314
? FILING DATE: 22-MAY-1995
? CLASSIFICATION: 435

```

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/170558
? FILING DATE: 20-DEC-1993
? PRIOR APPLICATION DATA: 08/157563
? FILING DATE: 23-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Haasak, Janet E.
? REGISTRATION NUMBER: 28,616
? REFERENCE/DOCKET NUMBER: 854C1D2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415/225-1896
? TELEFAX: 415/952-9881
? TELETYPE: 910/371-7168
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 913 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
?
? US-08-447-314-4
?
? Query Match 44.8%; Score 39; DB 3; Length 913;
? Best Local Similarity 38.5%; Pred. No. 4,4e+02;
? Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
?
? QY 3 MVLVPLILNGYTW 15
? : : : : : : : : : : : : : : : : : :
? DB 431 LLLIALMLRLHML 443
?
? RESULT 16
? US-08-445-461-4
? Sequence 4, Application US/08445461
? GENERAL INFORMATION:
? PATENT NO. 6096527
? APPLICANT: Mark, Melanie R.
? APPLICANT: Scadden, David T.
? APPLICANT: Baker, Kevin P.
? APPLICANT: Barton, Will F.
? TITLE OF INVENTION: Protein Tyrosine Kinases
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESS: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? PRIORITY DATE: 23-NOV-1993
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Palin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/445,461
? FILING DATE: 22-MAY-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/170558
? FILING DATE: 20-DEC-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/157563
? FILING DATE: 23-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Haasak, Janet E.
? REGISTRATION NUMBER: 28,616
? REFERENCE/DOCKET NUMBER: 854C3
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415/225-1896
? TELEFAX: 415/952-9881

```

? INFORMATION FOR SEQ ID NO: 4:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 913 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? US-08-445-461-4

Query Match 44.8% Score 39; DB 1; Length 913;
 Best Local Similarity 38.5% Pred. No. 4.4e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 3 MLVFFLLMGWTW 15
 DB 431 LLIATLMRLRM 443

RESULT 17
 US-08-336-343A-2
 ? Sequence 2, Application US/08336343A

? PATENT NO. 5695766
 ? GENERAL INFORMATION:
 ? APPLICANT: Ultrich, Axel
 ? APPLICANT: Alves, Frauke
 ? TITLE OF INVENTION: CCK-2, A No. 5677144e1 Receptor Tyrosine Kinase
 ? NUMBER OF SEQUENCES: 43
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Peunle & Edmunds
 ? ADDRESS: 1155 Avenue of the Americas
 ? CITY: New York
 ? STATE: New York
 ? COUNTRY: U.S.A.
 ? ZIP: 10036-2711
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: floppy disk
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30
 ? CURRENT APPLICATION NUMBER: US/08/336.343A
 ? FILING DATE: 08-NOV-1994
 ? CLASSIFICATION: 435
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Grogan, Laura A.
 ? REFERENCE NUMBER: 10.742
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (212) 790-9090
 ? TELEFAX: (212) 869-9741/8864
 ? TELEX: 66141 PENNIE
 ? INFORMATION FOR SEQ ID NO: 2:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1786 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: unknown
 ? MOLECULE TYPE: protein
 ? US-08-336-343A-2

Query Match 44.8% Score 39; DB 1; Length 919;
 Best Local Similarity 38.5% Pred. No. 4.4e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 3 MLVFFLLMGWTW 15
 DB 431 LLIATLMRLRM 443

RESULT 18
 US-08-477-451-16
 ? Sequence 16, Application US/08477451
 ? Patent No. 592885
 ? GENERAL INFORMATION:

? APPLICANT: Covacci, Antonello
 ? TITLE OF INVENTION: Helicobacter Pylori Cagl Region
 ? NUMBER OF SEQUENCES: 46
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Chiron Corporation
 ? STREET: 4560 Horton Street
 ? CITY: Emeryville
 ? STATE: CA
 ? COUNTRY: USA
 ? ZIP: 94608-2916
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: floppy disk
 ? OPERATING SYSTEM: IBM PC compatible
 ? SOFTWARE: PatentIn Release #1.0, Version #1.30
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/477.451
 ? FILING DATE: 07-JUN-1995
 ? CLASSIFICATION: 435
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: McClung, Herbert G.
 ? REFERENCE NUMBER: 10.113
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 510-601-2708
 ? TELEFAX: 510-655-3542
 ? INFORMATION FOR SEQ ID NO: 16:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1786 amino acids
 ? TYPE: amino acid
 ? STRANDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: protein
 ? US-08-477-451-16

Query Match 44.8% Score 39; DB 2; Length 1786;
 Best Local Similarity 40.8% Pred. No. 8.4e+02;
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Oy 1 MSLVFFLLMGWTW 15
 DB 807 MRLVFFPMQATW 821

RESULT 19
 US-08-131-625B-14
 ? Sequence 14, Application US/08131625B

? PATENT NO. 5695766
 ? GENERAL INFORMATION:
 ? APPLICANT: PAUL, PREM S.
 ? APPLICANT: HALBUR, PATRICK G.
 ? APPLICANT: WINK, JAMES-JIN
 ? APPLICANT: LITVO, MELISSA A.
 ? APPLICANT: LITVO, YOUNG S.
 ? TITLE OF INVENTION: VACCINES RAISING AN IMMUNOLOGICAL
 ? TITLE OF INVENTION: RESPONSE AGAINST VIRUSES CAUSING PORCINE RESPIRATORY AND
 ? TITLE OF INVENTION: REPRODUCTIVE DISEASES, METHODS OF PROTECTING A PIG AGAINST
 ? TITLE OF INVENTION: A DISEASE CAUSED BY A RESPIRATORY AND REPRODUCTIVE
 ? NUMBER OF SEQUENCES: 31
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: OHLSON, STYVAK, MCCLELLAND, MAIER & NEUBART,
 ? ADDRESS: P.C.
 ? STREET: 1755 S. Jefferson Davis Highway, Suite 400
 ? CITY: Arlington
 ? STATE: Virginia
 ? COUNTRY: U.S.A.
 ? ZIP: 22202
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: floppy disk
 ? OPERATING SYSTEM: IBM PC compatible
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/131.6258
 FILING DATE: 05-03-1993
 PRIORITY DATE: 05-03-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/969,071
 FILING DATE: 30-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 4625-016-55X CIP
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-131-6258-14

Query Match 44.3%, Score 38.5; DB 1; Length 174;
 Best Local Similarity 45.0%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 4; Indels 5; Gaps 1;
 Oy 1 MSMLVFLLMGV---TW 15
 Db 73 LTMGAVALMGVSALETW 92

RESULT 20
 US-08-799-464A-11
 Sequence 11; Application US/08799464A
 Patent No. 5998601
 GENERAL INFORMATION:
 APPLICANT: Murlaugh, Michael P. et al.
 TITLE OF INVENTION: VR-233 VIRAL NUCLEOTIDE SEQUENCE AND
 METHOD OF IDENTIFICATION
 ADDRESS: 1100
 CORRESPONDENCE ADDRESS:
 ADDRESSER: John M. Collins
 STREET: 2405 Grand Blvd., Suite 400
 CITY: Kansas City
 STATE: Missouri
 COUNTRY: USA
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/799,464A
 FILING DATE: 08/28/94
 PRIOR APPLICATION DATA:
 CLASSIFICATION: 536
 APPLICATION NUMBER: 08/287,941
 FILING DATE: August 5, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Collins, John M.
 REGISTRATION NUMBER: 6,263
 REFERENCE/DOCKET NUMBER: 4625-016-55X CIP
 TELEPHONE: (816) 474-9050
 TELEFAX: (816) 474-9057
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-799-464A-11

Query Match 44.3%, Score 38.5; DB 2; Length 174;
 Best Local Similarity 45.0%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 4; Indels 5; Gaps 1;
 Oy 1 MSMLVFLLMGV---TW 15
 Db 73 LTMGAVALMGVSALETW 92

RESULT 21
 US-08-686-968C-10
 Sequence 10; Application US/08686968C
 Patent No. 6221361
 GENERAL INFORMATION:
 APPLICANT: Cochran, Mark D.
 TITLE OF INVENTION: Recombinant Swinepox Virus
 FILE REFERENCE: 39119-H/JWL
 CURRENT APPLICATION NUMBER: US/08/686,968C
 CURRENT FILING DATE: 1996-07-25
 NUMBER OF SEQ ID NOS: 231
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 10
 TYPE: prt
 ORGANISM: Porcine reproductive and respiratory syndrome virus
 US-08-686-968C-10

Query Match 44.3%, Score 38.5; DB 4; Length 174;
 Best Local Similarity 45.0%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 4; Indels 5; Gaps 1;
 Oy 1 MSMLVFLLMGV---TW 15
 Db 73 LTMGAVALMGVSALETW 92

RESULT 22
 US-09-113-750A-41
 Sequence 41; Application US/09113750A
 Patent No. 6294176
 GENERAL INFORMATION:
 APPLICANT: David E. Junker and Mark D. Cochran
 TITLE OF INVENTION: Recombinant Raccoonpox Virus
 ADDRESS: 1100
 CORRESPONDENCE ADDRESS:
 ADDRESSER: John P. White
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/113,750A
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 55744
 TELEPHONE: (212)262-0400
 TELEFAX: (212)664-0525

TELEX: 422523
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-113-7504-41

Query Match 44.3% Score 38.5 DB 4 Length 174:
 Best Local Similarity 45.0% Pred. No. 1e+02:
 Matches 9: Conservative 2: Mismatches 4: Indels 5: Gaps 1:

1 MSMLVFLINCV-----TW 15
 DB 73 LTMCAVVALMGVSAIEFW 92

RESULT 23
 PCT-US95-09927-11
 Sequence 11, Application PC/TUS9509927
 GENERAL INFORMATION:
 APPLICANT: MURCAUGH, Michael P.
 TITLE OF INVENTION: VA-2332 VIRAL NUCLEOTIDE SEQUENCE AND
 TITLE OF INVENTION: METHODS OF USE
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESSES:
 ADDRESS: JONHAY COLLINS
 STREET: 1101 Walnut, Suite 1400
 CITY: Kansas City
 STATE: Missouri
 COUNTRY: USA
 ZIP: 64106
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentia Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/09927
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: COLLINS JOHN A.
 REGISTRATION NUMBER: 26122
 REFERENCE/DOCKET NUMBER: 22907
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (816) 474-9050
 TELEFAX: (816) 474-9057
 INFORMATION FOR SEQ ID NO: 11:
 TYPE: amino acid
 LENGTH: 174 amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-09927-11

Query Match 44.3% Score 38.5 DB 5 Length 174:
 Best Local Similarity 45.0% Pred. No. 1e+02:
 Matches 9: Conservative 2: Mismatches 4: Indels 5: Gaps 1:

1 MSMLVFLINCV-----TW 15
 DB 73 LTMCAVVALMGVSAIEFW 92

RESULT 24
 PCT-US95-10904-17
 Sequence 17, Application PC/TUS9510904
 GENERAL INFORMATION:

APPLICANT: PAUL, PREM S.
 APPLICANT: MENG, XIANG-JIN
 APPLICANT: HALBUR, PATRICK G.
 APPLICANT: MOROZOV, IGOR
 APPLICANT: LUN, MELISSA A.
 TITLE OF INVENTION: POLYNUCLEOTIC ACID ISOLATED FROM A
 TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV)
 TITLE OF INVENTION: A PROTEIN ENCODED BY THE POLYNUCLEOTIC ACID, A VACCINE
 TITLE OF INVENTION: PREPARED FROM OR CONTAINING THE POLYNUCLEOTIC ACID OR
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESSES:
 ADDRESS: OBLON, SHIVAK, MCCLELLAND, MAIER & NEUSTADT,
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentia Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10904
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M.P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 4625-021-55X CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-1200
 TELEFAX: (703) 413-2240
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-10904-17

Query Match 44.3% Score 38.5 DB 5 Length 174:
 Best Local Similarity 45.0% Pred. No. 1e+02:
 Matches 9: Conservative 2: Mismatches 4: Indels 5: Gaps 1:

1 MSMLVFLINCV-----TW 15
 DB 73 LTMCAVVALMGVSAIEFW 92

RESULT 25
 PCT-US95-10904-43
 Sequence 43, Application PC/TUS9510904
 GENERAL INFORMATION:
 APPLICANT: PAUL, PREM S.
 APPLICANT: MENG, XIANG-JIN
 APPLICANT: HALBUR, PATRICK G.
 APPLICANT: MOROZOV, IGOR
 APPLICANT: LUN, MELISSA A.
 TITLE OF INVENTION: A POLYNUCLEOTIC ACID ISOLATED FROM A
 TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV)
 TITLE OF INVENTION: A PROTEIN ENCODED BY THE POLYNUCLEOTIC ACID, A VACCINE
 TITLE OF INVENTION: PREPARED FROM OR CONTAINING THE POLYNUCLEOTIC ACID OR
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESSES:

ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10904

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/131,625
 FILING DATE: 05-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 4625-021-55X CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELETYPE: 248855 OPAT UR
 INFORMATION FOR SEO ID NO: 43:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-10904-43

Query Match 44.3% Score 38.5; DB 5; Length 174;
 Best Local Similarity 45.0%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

QY 1 MSKLVFLMGCV-----TW 15
 DB 73 LTMGAVALLMGVSALETW 92

RESULT 26
 PCT-US95-10904-45

GENERAL INFORMATION:
 SEQUENCE 45, Application PC/TUS9510904
 GENERAL INFORMATION: PREM S.
 APPLICANT: MENG, XIANG-JIN
 APPLICANT: HALBIR, PATRICK G.
 APPLICANT: KOROZOV, IGOR
 APPLICANT: LIM, MELISSA A.
 TITLE OF INVENTION: A POLYNUCLEIC ACID ISOLATED FROM A
 TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV),
 TITLE OF INVENTION: A PROTEIN ENCODED BY THE POLYNUCLEIC ACID, A VACCINE
 TITLE OF INVENTION: PREPARED FROM OR CONTAINING THE POLYNUCLEIC ACID OR
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSER: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10904

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/131,625
 FILING DATE: 05-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 4625-021-55X CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELETYPE: 248855 OPAT UR
 INFORMATION FOR SEO ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 174 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-10904-45

Query Match 44.3% Score 38.5; DB 5; Length 174;
 Best Local Similarity 45.0%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

QY 1 MSKLVFLMGCV-----TW 15
 DB 73 LTMGAVALLMGVSALETW 92

RESULT 27
 PCT-US95-10904-47

GENERAL INFORMATION:
 SEQUENCE 47, Application PC/TUS9510904
 GENERAL INFORMATION: PAUL, PREM S.
 APPLICANT: MENG, XIANG-JIN
 APPLICANT: HALBIR, PATRICK G.
 APPLICANT: KOROZOV, IGOR
 APPLICANT: LIM, MELISSA A.
 TITLE OF INVENTION: A POLYNUCLEIC ACID ISOLATED FROM A
 TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV),
 TITLE OF INVENTION: A PROTEIN ENCODED BY THE POLYNUCLEIC ACID, A VACCINE
 TITLE OF INVENTION: PREPARED FROM OR CONTAINING THE POLYNUCLEIC ACID OR
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSER: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10904

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/131,625
 FILING DATE: 05-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Lavalleye, Jean-Paul M. P.
 REGISTRATION NUMBER: 31,451
 REFERENCE/DOCKET NUMBER: 4625-021-55X CIP
 TELECOMMUNICATION INFORMATION:


```

? Patent No. 6110467
? GENERAL INFORMATION:
? APPLICANT: FIRM, PAUL S.
? INVENTOR: FIRM, PAUL S.
? INVENTOR: MENG, XIANG-JIN
? LYOO, MELISSA A.
? LYOO, YOUNG S.
? TITLE OF INVENTION: VACCINES RAISING AN IMMUNOLOGICAL
? RESPONSE AGAINST VIRUSES CAUSING PORCINE RESPIRATORY AND
? REPRODUCTIVE DISEASES, METHODS OF PROTECTING A PIG AGAINST
? A DISEASE CAUSED BY A RESPIRATORY AND REPRODUCTIVE
? NUMBER OF SEQUENCES: 47
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? P.C.
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
? CITY: ARLINGTON
? STATE: VIRGINIA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA: US/08/855,531D
? FILING DATE: 13-MAY-1997
? CLASSIFICATION: <unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/969,071
? FILING DATE: 30-OCT-1992
? ATTORNEY/AGENT INFORMATION: EMM, M.P.
? REGISTRATION NUMBER: 31,451
? REFERENCE/DOCKET NUMBER: 4625-038-55X DIV
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SEQ ID NO: 37:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 188 amino acids
? TYPE: amino acid
? STRANDEDNESS: linear
? TOPOLOGY: linear
? MOLECULE TYPE: Protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 37:
?
US-08-855-531D-37
?
Query Match 44.3%, Score 38.5; DB 3; Length 188;
Best Local Similarity 45.0%; Pied. No. 1.1e+02;
Matches 9; Conservative 4; Indels 5; Gaps 1;
?
QY 1 MSMLVYLLMGV----TW 15
DB 87 LTMCAVVALMGVSALETM 106
?
RESULT 31
US-08-855-526B-37
? Sequence 37, Application US/0885526B
? Patent No. 6251404
? GENERAL INFORMATION:
? APPLICANT: FIRM, PAUL S.
? INVENTOR: FIRM, PAUL S.
? INVENTOR: MENG, XIANG-JIN
? LYOO, MELISSA A.
? APPLICANT: LYOO, YOUNG S.
? TITLE OF INVENTION: VACCINES RAISING AN IMMUNOLOGICAL
? RESPONSE AGAINST VIRUSES CAUSING PORCINE RESPIRATORY AND
? REPRODUCTIVE DISEASES, METHODS OF PROTECTING A PIG AGAINST
? TITLE OF INVENTION: A DISEASE CAUSED BY A RESPIRATORY AND REPRODUCTIVE
? TITLE OF INVENTION: A DISEASE CAUSED BY A RESPIRATORY AND REPRODUCTIVE

```

```

? CORRESPONDENCE ADDRESS:
? ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
? CITY: ARLINGTON
? STATE: VA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/855,526B
? FILING DATE: 13-MAY-1997
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/969,071
? FILING DATE: 30-OCT-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: LAVALLIE, JEAN-PAUL M. P.
? REGISTRATION NUMBER: 31,451
? REFERENCE/DOCKET NUMBER: 4625-040-55X DIV
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SEQ ID NO: 37:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 188 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? MOLECULE TYPE: linear
? MOLECULE TYPE: peptide
?
US-08-855-526B-37
?
Query Match 44.3%, Score 38.5; DB 4; Length 188;
Best Local Similarity 45.0%; Pied. No. 1.1e+02;
Matches 9; Conservative 4; Indels 5; Gaps 1;
?
QY 1 MSMLVYLLMGV----TW 15
DB 87 LTMCAVVALMGVSALETM 106
?
RESULT 32
US-08-289-699A-4
? Sequence 4, Application US/08289699A
? Patent No. 5605993
? GENERAL INFORMATION:
? APPLICANT: Fukudome, Kenji
? APPLICANT: Esmon, Charles T.
? TITLE OF INVENTION: Cloning and Regulation of an Endothelial
? CELL PROTEIN C/ACTIVATED PROTEIN C RECEPTOR
? NUMBER OF SEQUENCES: 6
? SOURCE: HUMAN
? ADDRESS:
? ADDRESSEE: Palmetto, Robert
? STREET: 2800 One Atlantic Center, 1201 West Peachtree
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: US
? ZIP: 30345
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/289,699A

```

```

? FILING DATE: 12-AUG-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.
? REGISTRATION NUMBER: 31,284
? TELEPHONE: (404)873-8794
? TELEFAX: (404)873-8794
? TELECOMMUNICATION INFORMATION:
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 335 amino acids
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
?
US-08-289-6394-4

Query Match 43.7%; Score 38; DB 1; Length 335;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Oy 5 VYFLLIAGV--TWG 16
Db 4 LFLILLALLDAMG 17

RESULT 33
US-08-878-283-4
? Sequence 4, Application US/08878283
? Patent No. 5852171
? GENERAL INFORMATION:
? APPLICANT: Fukudome, Kenji
? APPLICANT: Esmon, Charles T.
? TITLE OF INVENTION: Cloning and Regulation of an Endothelial
? TITLE OF INVENTION: C11 Protein C/Activated Protein C Receptor
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Patrea L. Pabst
? STREET: 2800 One Atlantic Center, 1201 West Peachtree
? STREET: Street
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: US
? ZIP: 30306-3450
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? FILING DATE: US/08/878,283
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/289,699
? FILING DATE: 12-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.
? REGISTRATION NUMBER: 31,284
? TELEPHONE: (404)873-8794
? TELEFAX: (404)873-8794
? TELECOMMUNICATION INFORMATION:
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 335 amino acids
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
?
US-08-878-283-4

```

```

Query Match 43.7%; Score 38; DB 2; Length 335;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Oy 5 VYFLLIAGV--TWG 16
Db 4 LFLILLALLDAMG 17

RESULT 34
US-08-852-824-4
? Sequence 15, Application US/08852824C
? Patent No. 6060272
? GENERAL INFORMATION:
? APPLICANT: Li, et al.
? TITLE OF INVENTION: Human G-Protein Coupled Receptors
? FILE REFERENCE: 1488,1220000
? CURRENT APPLICATION NUMBER: US/08/852,824C
? CURRENT FILING DATE: 1997-05-04
? NUMBER OF SEQUENCES: 18
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 4
? LENGTH: 384
? TYPE: PRT
? ORGANISM: genomic
?
US-08-852-824-4

Query Match 43.7%; Score 38; DB 3; Length 384;
Best Local Similarity 37.5%; Pred. No. 2.6e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 1 MSNIVFLIAGV--TWG 16
Db 254 LMLTLLIAGV--TWG 269

RESULT 35
US-08-385-186-15
? Sequence 15, Application US/08385186
? Patent No. 5744594
? GENERAL INFORMATION:
? APPLICANT: Ashford, John P.
? APPLICANT: Bond, Chris T.
? TITLE OF INVENTION: DNA ENCODING ATP-SENSITIVE POTASSIUM
? TITLE OF INVENTION: CHANNEL PROTEINS AND USES THEREOF
? NUMBER OF SEQUENCES: 15
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Townsend and Townsend Kourile and Crew
? STREET: 1500 California Street, Tower, One Market Plaza
? CITY: San Francisco
? STATE: California
? COUNTRY: US
? ZIP: 94105-1493
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/385,186
? FILING DATE: 07-FEB-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/288,510
? FILING DATE: 10-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.
? REGISTRATION NUMBER: 31,284
? TELEPHONE: (404)873-8794
? TELEFAX: (404)873-8794
? TELECOMMUNICATION INFORMATION:
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 384 amino acids
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
?
US-08-385-186-15

```



```

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED FOR Windows Version 2.0
CROSS-REFERENCE NUMBER: US/08/790,374
APPLICATION NUMBER: US/08/845,566
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
ADDRESS: 156, 749
REFERENCE/DOCKET NUMBER: PF-0271 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TYNMOR01
CLONE: 144690
US-08-845-566-1

Query Match 43.7%; Score 38; DB 2; Length 509;
Best Local Similarity 37.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 MSALVELLAWYNG 16
DB 254 LMLLALLVCMGFLGE 269

RESULT 39
US-08-790-374-2
Sequence 2, Application US/08790374
Patent No. 5863734
GENERAL INFORMATION:
APPLICANT: Koryolougru, Valtia
INVENTOR: Koryolougru, Valtia
TITLE OF INVENTION: METHODS OF TREATMENT FOR OBSESSIVE-COMPULSIVE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CROSS-REFERENCE NUMBER: US/08/790,374
APPLICATION NUMBER: US/08/790,374
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

```

```

NAME: Jackson, Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-189
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-467-5800
TELEFAX: 201-515-1664
TELEX: 133521
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY:
HYPOTHETICAL: NO
US-08-790-374-2

Query Match 43.1%; Score 37.5; DB 2; Length 271;
Best Local Similarity 58.8%; Pred. No. 2.1e+02;
Matches 10; Conservative 1; Mismatches 1; Indels 5; Gaps 2;

OY 3 MLYVELL---MCTYNG 16
DB 18 LLYVLLLLRHMC--MG 32

RESULT 40
PCT-US96/10602-2
Sequence 2, Application PC/TUS9610602
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 315 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10602
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017,814
CLASSIFICATION: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8506
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY:
HYPOTHETICAL: NO
PCT-US96/10602-2

Query Match 43.1%; Score 37.5; DB 5; Length 346;
Best Local Similarity 41.2%; Pred. No. 2.7e+02;

```

Matches 7; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
OY 1 MSNLYVFLIMGV-TWC 16
Db 306 ISLNVFLIMVIMVWG 322

RESULT 41
US-08-157-005-2
Sequence 2, Application US/08157005
Patent No. 5620691
GENERAL INFORMATION:
APPLICANT: Meusvoort, Gert
APPLICANT: Terpstra, Cathrinus
APPLICANT: Pol, Johannes M
APPLICANT: Moorman, Robertus J
APPLICANT: Meulenber, Johanna J
TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
STREET: 525 South 100 East
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,863
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/157,005
FILING DATE: 26-NOV-1999
PRIOR APPLICATION NUMBER: EP 91201398.4
FILING DATE: 06-JUN-1991
APPLICATION DATA:
APPLICATION NUMBER: EP 92200781.0
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00096
FILING DATE: 05-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Turner, Allen C.
REGISTRATION NUMBER: 33041
TELEPHONE: (801) 532-1992
TELEFAX: (801) 531-9168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2396 amino acids
TOPOLGY: 1 linear
MOLECULE TYPE: protein
US-08-157-005-2

Query Match 43.1%, Score 37.5; DB 1; Length 2396;
Best Local Similarity 44.4%; Pctd. No. 1,9e+03;
Matches 8; Conservative 2; Mismatches 5; Indels 3; Gaps 1;
OY 1 MSNLYVFLIMGV-TWC 15
Db 1902 VOLICVFLIMVIMVWG 1919

RESULT 42
US-08-747-863-2

Sequence 2, Application US/08747863
Patent No. 6197310
GENERAL INFORMATION:
APPLICANT: Meusvoort, Gert
APPLICANT: Terpstra, Cathrinus
APPLICANT: Pol, Johannes M
APPLICANT: Moorman, Robertus J
APPLICANT: Meulenber, Johanna J
TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
STREET: 525 South 100 East
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,863
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/157,005
FILING DATE: 26-NOV-1999
PRIOR APPLICATION NUMBER: EP 91201398.4
FILING DATE: 06-JUN-1991
APPLICATION DATA:
APPLICATION NUMBER: EP 92200781.0
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00096
FILING DATE: 05-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Turner, Allen C.
REGISTRATION NUMBER: 33041
TELEPHONE: (801) 532-1992
TELEFAX: (801) 531-9168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2396 amino acids
TOPOLGY: 1 linear
MOLECULE TYPE: protein
US-08-747-863-2

Query Match 43.1%, Score 37.5; DB 4; Length 2396;
Best Local Similarity 44.4%; Pctd. No. 1,9e+03;
Matches 8; Conservative 2; Mismatches 5; Indels 3; Gaps 1;
OY 1 MSNLYVFLIMGV-TWC 15
Db 1902 VOLICVFLIMVIMVWG 1919

RESULT 43
PCTUS92-00282-9
Sequence 9, Application PC/TUS9200282
GENERAL INFORMATION:
APPLICANT: Owens, Ida S.
APPLICANT: Ritter, Joseph K.
TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
THEREIN.
NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 STREET: 1615 L STREET, N.W.
 CITY: WASHINGTON
 STATE: DISTRICT OF COLUMBIA
 COUNTRY: U.S.A.
 ZIP: 20036-5601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CUSHMAN DARBY & CUSHMAN
 APPLICATION NUMBER: PCT/US92/00282
 FILING DATE: 19920110
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: SCOTT, WATSON T.
 REGISTRATION NUMBER: 26581
 REFERENCE/DOCKET NUMBER: 91532-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-661-3000
 TELEFAX: 202-822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 286 amino acids
 TYPE: AMINO ACID
 STRANDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US92-00282-9

Query Match 42.5% Score 37; DB 5; Length 286;
 Best Local Similarity 38.3%; Pred. NO. 2,6e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 5 LVFLFMVYTW 16
 DB 15 VFFLALMGVYG 26

RESULT 44
 US-08-385-186-2
 Sequence 2: Application US/08385186
 Patent No. 5744594
 GENERAL INFORMATION:
 APPLICANT: Adelman, John P
 APPLICANT: Ashford, Michael J
 TITLE OF INVENTION: DNA ENCODING ATP-SENSITIVE POTASSIUM
 CHANNEL PROTEINS AND USES THEREOF
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourile and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94105-1493
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/385,186
 FILING DATE: 07-FEB-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/288,510
 FILING DATE: 10-AUG-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/193,372
 FILING DATE: 08-FEB-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: SCOTT, WATSON T.
 REGISTRATION NUMBER: 26581
 REFERENCE/DOCKET NUMBER: 91532-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 494 amino acids
 TYPE: AMINO ACID
 STRANDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-385-186-2

Query Match 42.5% Score 37; DB 1; Length 419;
 Best Local Similarity 38.3%; Pred. NO. 3,3e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 4 LVFLFMVYTW 15
 DB 90 LVFLFMVYTW 101

RESULT 45
 US-09-006-636-8
 Sequence 8: Application US/0906636
 Patent No. 6005092
 GENERAL INFORMATION:
 APPLICANT: Shoseyov, Oded
 APPLICANT: Shani, Ziv
 TITLE OF INVENTION: ADRIATOPIS THALIANA ENDO-1,4-BETA-
 D-GALACTOSYLTRANSFERASE GENE, PROMOTER AND PROTEIN
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penile & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10036-5711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/006,636
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Baldwin, Geraldine F.
 REGISTRATION NUMBER: 31,232
 REFERENCE/DOCKET NUMBER: 7809-019
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 869-8900
 TELEFAX: (212) 869-8664
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 494 amino acids
 TYPE: amino acid
 STRANDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-006-636-8

Query Match 42.5% Score 37; DB 3; Length 494;

Best Local Similarity 71.4%; Prod. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels

0; Gaps 0;

Oy 9 LLMCWTM 15
111111
DB 241 LLMCASH 247

Search completed: January 7, 2002, 16:49:57
Job time: 284 sec

A:Residues: 1-298 <PAR>
 A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PID:CA873960.1; PID:9696896
 A:Experimental source: serotype O2, strain NCTC 11168
 A:Gene: CJ1544c

Query Match 50.6%: Score 44; DB 2; Length 298;
 Best Local Similarity 46.2%: Pred. No. 36;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 3 MVVFLLMGVTW 15
 Db 10 LVIVLWFLMGSSW 22

RESULT 19
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #excl_change 21-Jan-2000
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Accession: T33734
 R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: 221588
 A:Accession: T33734
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <SMD>
 A:Cross-references: EMBL:AL109732; PID:CA852045.1; GSPDB:GN00070; SCEDB:SC7H2_03c
 A:Experimental source: strain A3(2)
 A:Gene: SC7H2_03c
 A:Superfamily: proline aminopeptidase

Query Match 50.6%: Score 44; DB 2; Length 323;
 Best Local Similarity 75.0%: Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 9 LLMGVTW 16
 Db 109 LVIVLWFLMGSSW 116

RESULT 20
 T04608
 ADP-ATP carrier protein P209.60 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 22-Apr-1999 #sequence_revision 23-Apr-1999 #excl_change 21-May-1999
 R:Brown, M.; Rose, M.; Hompl, S.; Entlin, K.D.; Hohnselt, J.; Mewes, H.W.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, October 1998
 A:Reference number: 215380
 A:Accession: T04608
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-379 <BEV>
 A:Cross-references: EMBL:AL021749
 A:Experimental source: cultivar Columbia; BAC clone P2095
 C:Gene: ATP5C1
 A:Map position: 4
 A:Introns: 157/3; 281/3
 A:Note: P209.60
 C:Superfamily: ADP-ATP carrier protein; ADP-ATP carrier protein repeat homology
 C:Keywords: duplication; transmembrane protein
 F:6-112/Domain: ADP-ATP carrier protein repeat homology <AC13>
 F:16-172/Domain: ADP-ATP carrier protein repeat homology <AC15>
 F:281-369/Domain: ADP-ATP carrier protein repeat homology <AC15>

Query Match 50.6%: Score 44; DB 2; Length 379;
 Best Local Similarity 60.0%: Pred. No. 43;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 2 SMVFLLMGVTW 16
 Db 283 SFVSLFLMGVTIC 297

RESULT 21
 H83708
 hypothetical protein BH0472 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #excl_change 31-Dec-2000
 R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: 221588
 A:Accession: H83708
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <SMD>
 A:Cross-references: GB:AP001508; NID:g10172890; PID:BA804191.1; GSPDB:C
 A:Experimental source: strain C-125
 A:Gene: BH0472

Query Match 50.6%: Score 44; DB 2; Length 417;
 Best Local Similarity 54.5%: Pred. No. 48;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 5 VVFLLMGVTW 15
 Db 104 VLRLLMGIAW 114

RESULT 22
 T49017
 hypothetical protein F3C22.40 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #excl_change 18-Aug-2000
 C:Accession: T49017
 R:Purcell, B.; Masuy, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 425013
 A:Accession: T49017
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-486 <PVR>
 A:Cross-references: EMBL:AL353912; GSPDB:GN00061; ANSP:F3C22.40
 A:Experimental source: cultivar Columbia; BAC clone F3C22
 A:Gene: ATP5C1
 A:Map position: 3
 A:Introns: 18/3; 43/2; 87/3; 111/3; 191/3; 175/3; 222/3; 251/1; 399/2
 C:Superfamily: Arabidopsis thaliana hypothetical protein F3C22.40

Query Match 50.6%: Score 44; DB 2; Length 486;
 Best Local Similarity 66.7%: Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 5 VVFLLMGVTW 16
 Db 112 LVVFLWFLMGSSW 123

RESULT 23
 H96681
 protein P1R22.10 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #excl_change 31-Mar-2001
 C:Accession: H96681
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

C:Species: *Neisseria meningitidis*
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #texL_change 02-Feb-2001
 C:Accession: AB1906
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 M.H.; Young, P.; Parkhill, J.; Leach, S.; Moule, S.; Kung'u, M.A.; Rajandream,
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.
 A:Reference number: AB1775; MUID:2022556
 A:Accession: AB1906
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1,218 <K>K>
 A:Experimental source: serogroup A, strain 22491
 C:Genetics:
 A:Gene: NMA1368

Query Match 48.9% Score 42.5; DB 2; Length 218;
 Best Local Similarity 41.5% Pred. No. 45;
 Matches 0; Conservative 2; Mismatches 1; Gaps 1;
 Oy 4 LWVFLWCVYWG 16
 Db 50 LVRRLCWGI-WG 61

RESULT 37
 hypothetical protein 2 - *Escherichia coli* plasmid ColD
 C:Species: *Escherichia coli*
 C:Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #texL_change 18-Jun-1993
 C:Accession: J25035
 R:Manukovich, J.A.; Han, C.H.; Konisky, J.
 A:Title: DNA and amino acid sequence analysis of structural and immunity genes of colic
 A:Reference number: A91822; MUID:87008385
 A:Accession: J25035
 A:Molecule type: DNA
 A:Residues: 1,102 <K>K>
 C:Genetics:
 A:Gene: plasmid

Query Match 48.3% Score 42; DB 2; Length 102;
 Best Local Similarity 71.4% Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 0; Gaps 0;
 Oy 10 LWCVYWG 16
 Db 44 LWGSMG 50

RESULT 38
 hypothetical protein 2 - *Escherichia coli* plasmid ColD
 C:Species: *Escherichia coli*
 C:Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #texL_change 18-Jun-1993
 C:Accession: J25035
 R:Manukovich, J.A.; Han, C.H.; Konisky, J.
 A:Title: DNA and amino acid sequence analysis of structural and immunity genes of colic
 A:Reference number: A91822; MUID:87008385
 A:Accession: J25035
 A:Molecule type: DNA
 A:Residues: 1,102 <K>K>
 C:Genetics:
 A:Gene: plasmid

Query Match 48.3% Score 42; DB 2; Length 102;
 Best Local Similarity 71.4% Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 0; Gaps 0;
 Oy 10 LWCVYWG 16
 Db 44 LWGSMG 50

Oy 10 LWCVYWG 16
 Db 44 LWGSMG 50

RESULT 39
 Probable nucleotide ATP-binding protein I APE1548 - *Aeropyrum pernix* (strain K1)
 C:Species: *Aeropyrum pernix*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #texL_change 17-Mar-2000
 C:Accession: E72636
 R:Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatake, Y.; Uda, K.; Ta
 DNA Res: 6,83,101,1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Cremarchaeon*, *Aero*
 A:Reference number: A72450; MUID:99310339
 A:Accession: E72636
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1,338 <K>K>
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1548
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homol

Query Match 48.3% Score 42; DB 2; Length 338;
 Best Local Similarity 57.1% Pred. No. 74;
 Matches 8; Conservative 1; Mismatches 5; Gaps 0;
 Oy 3 MLVFLWCVYWG 16
 Db 25 MLGCVYWG 38

RESULT 40
 hypothetical protein BH1161 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #texL_change 31-Dec-2000
 C:Accession: AB379596
 R:Yasuda, K.; Takai, Y.; Meno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res 28:4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
 A:Reference number: AB3650; MUID:20263314
 A:Accession: AB3795
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1,338 <K>K>
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1161

Query Match 48.3% Score 42; DB 2; Length 406;
 Best Local Similarity 57.0% Pred. No. 65;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Oy 4 LWVFLWCVYWG 15
 Db 97 LVRRLCWGI-WG 108

RESULT 41
 Probable integral membrane transport protein - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #texL_change 21-Jan-2000
 C:Accession: T36979
 R:Oliver, K.; Horlitz, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A

submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21618
 A:Accession: J15679
 A:Status: preliminary; translated from GR/EMBL/DBJ
 A:Key words: octopus rhodopsin
 A:Residues: 1-407
 A:Cross-references: EMBL:AL109949; PIDN:CA52893.1; GSPDB:ON00070; SCODEB:SCJ11_08C
 A:Genetics:
 C:gene: SCODha:SCJ11_08c
 C:superfamily: Streptomyces lividans chloramphenicol resistance protein

Query Match 48.3% Score 42; DB 2; Length 407;
 Best Local Similarity 58.3%; Pred. No. 86;
 Matches 7; Conservative 2; Mismatches 0; Gaps 0;
 Indels 0;

Y 5 MVFLMGVTCG 16
 DB 313 VALVFLMGVTCG 324
 :::::|:::|:::|

RESULT 42
 S24055
 preprotein translocase secY [validated] - Methanococcus vannielii
 C:Species: Methanococcus vannielii
 C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 A:Accession: S24055 G; Boeck, A.
 R:Biochem J 320:683-688 1997
 A:Title: Presence of a gene in the archaeobacterium Methanococcus vannielii homologous to
 A:Reference number: S24055; MUID:92110434
 A:Accession: S24055
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 428
 A:Cross-references: EMBL:X62045; NID:944771; PIDN:CA43978.1; PID:944772
 C:Function:
 A:Description: essential for preprotein translocation across the cytoplasmic membrane [Y
 A:Note: Methanococcus secY protein can complement a ts mutant in Escherichia coli
 C:Superfamily: Yeast SSB1 protein
 C:Keywords: protein transport; transmembrane protein
 F:30-46/Domain: transmembrane #status predicted <TM>
 F:46-112/Domain: transmembrane #status predicted <TM>
 F:112-128/Domain: transmembrane #status predicted <TM>
 F:128-154/Domain: transmembrane #status predicted <TM>
 F:159-185/Domain: transmembrane #status predicted <TM>
 F:211-227/Domain: transmembrane #status predicted <TM>
 F:252-268/Domain: transmembrane #status predicted <TM>
 F:318-334/Domain: transmembrane #status predicted <TM>
 F:377-393/Domain: transmembrane #status predicted <TM>

Query Match 48.3% Score 42; DB 1; Length 438;
 Best Local Similarity 46.2%; Pred. No. 91;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 Indels 0;

Y 3 MCVFLMGVTCG 15
 DB 323 MCVFLMGVTCG 335
 :::::|:::|:::|

RESULT 43
 000CG
 Rhodopsin - giant octopus
 C:Species: Octopus joffeini (giant octopus)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 05-Jun-1998
 R:Ouchimkov, Y.A.; Abduleev, N.G.; Zojicic, A.S.; Artamonov, I.D.; Bepko, I.A.; I
 FRS Lett. 232: 69-77, 1988
 A:Title: Octopus rhodopsin. Amino acid sequence deduced from cDNA.
 A:Reference number: S00610; MUID:88211878
 A:Accession: S00610

A:Molecule type: mRNA
 A:Residues: 1-455
 A:Cross-references: EMBL:X07797
 A:Note: the source is designated as Pteroctopus defileini
 C:Superfamily: octopus rhodopsin
 C:Keywords: chromoprotein; color vision; G protein-coupled receptor; glycoprotein; 11
 F:37-61/Domain: transmembrane #status predicted <TM>
 F:74-98/Domain: transmembrane #status predicted <TM>
 F:107-131/Domain: transmembrane #status predicted <TM>
 F:153-177/Domain: transmembrane #status predicted <TM>
 F:201-224/Domain: transmembrane #status predicted <TM>
 F:261-284/Domain: transmembrane #status predicted <TM>
 F:302-323/Domain: transmembrane #status predicted <TM>
 F:9-15/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:106/Binding site: retinal (Lys) (covalent) #status predicted
 F:337,338/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 48.3% Score 42; DB 1; Length 455;
 Best Local Similarity 23.1%; Pred. No. 97;
 Matches 3; Conservative 7; Mismatches 0; Gaps 0;
 Indels 0;

Y 3 MCVFLMGVTCG 15
 DB 155 LMIFVFMGVTGW 167
 :::::|:::|:::|

RESULT 44
 S28306
 hypothetical protein T2365.5 - Caenorhabditis elegans (fragment)
 C:Species: Caenorhabditis elegans
 C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Aug-1998
 A:Accession: S28306
 R:Beck, M.
 Submitted to the EMBL Data Library, December 1992
 A:Reference number: S28306
 A:Accession: S28306
 A:Molecule type: DNA
 A:Residues: 1-499 <BB>
 A:Cross-references: EMBL:219158
 C:Genetics:
 A:Interons: 27/3: 134/2; 215/3: 253/2; 332/3: 375/3: 409/1; 442/3
 C:Superfamily: gamma-aminobutyric acid transporter

Query Match 48.3% Score 42; DB 2; Length 499;
 Best Local Similarity 35.7%; Pred. No. 18+02;
 Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 Indels 0;

Y 3 MCVFLMGVTCG 16
 DB 379 LNMCHMLITGTWG 392
 :::::|:::|:::|

RESULT 45
 G64902
 N-terminus acid resistance protein xaaS - Escherichia coli
 A:Alternate names: probable peptidase xaaS
 C:Species: Escherichia coli
 C:Date: 2-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
 C:Accession: G64902; S78627
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Berna, N.T.; Burland, V.; Riley, M.;
 Science 277: 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A67470; MUID:97426617
 A:Accession: G64902
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-511 <BLAT>
 A:Cross-references: GB:A000246; GB:U00099; NID:9187764; PIDN:AC74565.1; PID:918777
 A:Experimental source: strain K-12, substrain MG1655

R:herst, B.M.; Farooq, F.T.; Barstad, D.N.; Blankenhorn, D.L.; Slonczewski, J.L.
J. Bacteriol. 178, 3978-3981, 1996
A:Title: A glutamate-dependent acid resistance gene in *Escherichia coli*.
A:Reference number: S78627; MUID:96272279
A:Accession number: U013204
A:Molecule type: DNA
A:Status nucleic acid sequence not shown; translation not shown
A:Residues: 402-489 <HER>
A:CROSS-references: EMBL:U013204
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C:Gene(s): gacC
C:Gene(s): gacC
C:Description: may act as a transporter for the product of glutamate decarboxylase, gamma
A:Note: part of a glutamate decarboxylase alkalization cycle to protect *E. coli* from C
C:Superfamily: arginine permease
C:Keywords: amino acid transport; inner membrane; Transmembrane protein
F:13-29/Domain: transmembrane #status predicted <TM1>
F:32-58/Domain: transmembrane #status predicted <TM2>
F:62-113/Domain: transmembrane #status predicted <TM3>
F:128-144/Domain: transmembrane #status predicted <TM4>
F:161-177/Domain: transmembrane #status predicted <TM5>
F:201-217/Domain: transmembrane #status predicted <TM6>
F:236-252/Domain: transmembrane #status predicted <TM7>
F:290-306/Domain: transmembrane #status predicted <TM8>
F:339-355/Domain: transmembrane #status predicted <TM9>
F:373-389/Domain: transmembrane #status predicted <TM10>
F:413-429/Domain: transmembrane #status predicted <TM11>
F:448-464/Domain: transmembrane #status predicted <TM12>

Query Match 48.3% Score 42; DB 2; Length 511;
Best Local Similarity 66.7% Pred. No. 16+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Ox 4 LVFLLLCYTM 15
||| ||| | |
DB 42 LVFLLLCGLW 53

Search completed: January 7, 2002, 16:52:10
Job time: 227 sec


```

YPR_ECOLI STANDARD: PRT: 217 AA.
AC P13974:1890 (Rel. 13, Created)
DT 01-JUN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHEICAL 24.3 KDA PROTEIN (YRF 1.)
OS Escherichia coli.
OC Plasmid IncFII R1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Neisseriella.
CX NCBI_TaxID:562;
RP SEQUENCE FROM N.A.
RX MEDLINE:88121694; Pubmed:332833;
RX Bravo A., de Tornyegui G., Diaz R.;
RX "Identification of components of a new stability system of plasmid
RX R1, PARD, that is close to the origin of replication of this
RX plasmid."
RL MedGen: Secret. 210:101-110(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC or send an email to license@isb.slb.ch). See http://www.isb.slb.ch/announce/
CC -----
DR EMBL: X06240; CNA29583.1; -.
DR PIR: S01095; S01095.
DR InterPro: IP0003675; Ab1.
DR Pfam: PF02517; Ab1; 1.
DR EMBL: X06240:1-110(1987).
RX SEQUENCE 217 AA; 24307 MW; EA51991CA3b6666 CRC64;
SO SEQUENCE

Query Match 56.38; Score 49; DB 1; Length 217;
Best Local Similarity 37.58; Pred. No. 3.9;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
OY 1 MENVYFLLMGVING 16
DB 9 LQPLVHLMAMGISMG 24

RESULT 2
YPR_ECOLI STANDARD: PRT: 1107 AA.
AC P13285:976708;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL PROTEIN YPR PRECURSOR.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Neisseriella.
CX NCBI_TaxID:562;
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE:95334362; Pubmed:7610040;
RX Burland V.O., Plunkett G. III, Sofia H.J., Daniels D.L.,
RX Blattner F.R.;
RX "The complete genome sequence of Escherichia coli genome VI: DNA sequence of the
RX region from 92.8 through 100 minutes."
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [2]
RP REVISION TO 1015.
RX STRAIN-K12 / MG1655;
RX MEDLINE:97426617; Pubmed:9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Berna M.T., Burland V.,

```

```

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RX Science 277:1235-1244(1997).
RN [3]
RP SEQUENCE OF 1-27 FROM N.A.
RX STRAIN-K12;
RX MEDLINE:88298809; Pubmed:3042771;
RX Li O.X., Dowhan W.;
RX "Structural characterization of Escherichia coli phosphatidylserine
RX transferase YnfP."
RX J Biol Chem. 269:11516-11522(1998).
CC -1- SUBCELLULAR LOCATION: INTERAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE UPF0003 FAMILY. STRONG, TO H. INFLUENZAE
CC H10195.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb.slb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: U14003; AAB70591.1; -.
DR EMBL: X06240:881-110(1987); 1; AUT_UNIT.
DR EMBL: J03916; AAB3697.1; AUT_UNIT.
DR Ecocore: E012478; Y1ep.
DR InterPro: IP001880; UPF0003.
DR Pfam: PF00924; UPF0003; 1.
DR PROSITE: PS01246; UPF0003; 1.
DE Hypothetical protein; Transmembrane; Signal; Complete proteome.
OS SIGNAL.
CX NCBI_TaxID: 21099;
CX EMBL: X06240:881-110(1987); 1;
FT TRANSMEM 467 487 POTENTIAL.
FT TRANSMEM 551 571 POTENTIAL.
FT TRANSMEM 600 620 POTENTIAL.
FT TRANSMEM 628 648 POTENTIAL.
FT TRANSMEM 674 694 POTENTIAL.
FT TRANSMEM 698 718 POTENTIAL.
FT TRANSMEM 735 755 POTENTIAL.
FT TRANSMEM 825 845 POTENTIAL.
FT TRANSMEM 875 895 POTENTIAL.
FT TRANSMEM 910 930 POTENTIAL.
FT TRANSMEM 1015 1015 POTENTIAL.
FT CONFLICT 1015 1015 R > A (IN REF. 1).
SO SEQUENCE 1107 AA; 123967 MW; 5F5A2A93B90532B CRC64;

Query Match 56.38; Score 49; DB 1; Length 1107;
Best Local Similarity 37.58; Pred. No. 16;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
OY 1 MENVYFLLMGVING 16
DB 1 MRLTFVLMAMGISMG 16

RESULT 3
YC21_ARCFU STANDARD: PRT: 48 AA.
AC O29047;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL PROTEIN Ar1221.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
CX NCBI_TaxID:2234;
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-VG-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=90049343; PubMed=9384475;
RA Klenk H.-P., Dodson R.A., Tomb J.-F., White O., Nelson K.E.,
RA Richardson D., Dolan E.P., Besser T.E., Peterson S.J.D.,
RA Fleischmann R.D., Kurland A.R., Gnanapavan D.B., Kyriakidis
RA Kirchman R.F., Dougherty B.A., McKenney K., Adams M.D., Loftis B.,
RA Peterson S., Reich C.I., Mellé L.K., Badger J.H., Glöckner A., Zhou L.,
RA Overbeck R., Coagayne J.D., Weidman J.F., McDonald L., Utechback T.,
RA Cotton M.D., Spriggs T., Atleick P., Kalne B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Webster J.C., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RL reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The content of this entry has been derived from scientific publications and may
CC use by non-profit institutions as long as its content is not modified or
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC CC
CC EMBL: AE001020; MAF950036.1; .
DR TTIG: AE1221;
DR NCBI: AF1221;
DR TRANSFAM: T5247;
FT SEQUENCE 48 AA: 5307 MW: 60M65535B27BC1 CRC64:
SO
OY 5 VFFLLKMTWY 15
DB 37 IVILLINLSMW 47

RESULT_4
1006 BORBU STANDAO: PRT: 374 AA.
1006 BORBU
AC OS1039:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEETICAL PROTEIN BB0006.
GN BB0006.
OS Borrelia burgdorferi ( Lyme disease spirochete ).
OX NCBI_TaxId=139.
RX MEDLIN=98065943; PubMed=94033685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Fleischman R.A., Dodson R., Hickey E.K., Gwinn M.,
RA Peterson J., Kurland A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Coagayne J.D., Weidman J.,
RA Utechback T., Wattey L.T., McDonald L., Atleick P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RT Nature 390:565-566(1997).
CC - SIMILARITY: BELONGS TO THE UPDOLIS (PSRM) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.

```

CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is prominently displayed.
CC	Outstanding requirements for commercial use can be found at http://www.ebi.ac.uk/biomed/ .
CC	Send us your comments via email to license@ebi.ac.uk .
DR	EMBL; A601115; M066397.1; -
DR	TIGR; BB0006;
DR	InterPro; IPRO02549; UPF0118.
DR	Pfam; PF01594; UPF0118.1
KM	Hypothetical protein: Transmembrane; Complete proteome.
NW	TRANSMEM 52 75 POTENTIAL..
FT	TRANSMEM 52 75 POTENTIAL..
FT	TRANSMEM 80 100 POTENTIAL..
FT	TRANSMEM 157 177 POTENTIAL..
FT	TRANSMEM 224 244 POTENTIAL..
FT	TRANSMEM 246 266 POTENTIAL..
FT	TRANSMEM 310 330 POTENTIAL..
FT	TRANSMEM 336 356 POTENTIAL..
SQ	SEQUENCE 374 AA: 42820 MW: 9053960.4B0645900 CRC64;
OY	3 MYVELLUNGTCG 16
Dc	311 LILCFIFPNMHAK 324
RESULT	5
PIP_STRCO	STANDARD: PRT: 323 AA.
ID	PIP_STRCO
AC	QOSQLI: 2001 (Rel. 40, Created)
DT	20-ANG-2001 (Rel. 40, Last sequence update)
DI	20-ANG-2001 (Rel. 40, Last annotation update)
DT	PROBABE PROLINE ISOMERITIDASE (EC 3.4.11.5) (PIP) (PROLYL
GN	AMINOPEPTIDASE) (PAP).
DN	SC7H2_03C.
OS	Streptomyces coelicolor.
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetia;
NC	Streptomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX	Neisseria1905;
RX	[1]
RF	SEQUENCE FROM N.A.
RC	STRAIN-AJ(2);
RA	Sanders D.C., Harris D., Bentley S.D., Parkhill J., Barrell B.G.,
RL	Submitted (Aug-1999) to the EMBL/Genbank/DDJJ databases.
RU	-1- RESIDUES FROM PEPTIDES (BY SIMILARITY)
CC	+1 CATALYTIC ACTIVITY: RELEASE OF A N-TERMINAL PROLINE FROM A
CC	PEPTIDE.
CC	-1 SUBCELLULAR LOCATION: CYTOSOLASMIC (BY SIMILARITY).
CC	-1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S33.
CC	This SWISS-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Data Bank
CC	on the European Bioinformatics Institute's behalf.
CC	Use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isdb.ch/announce/) or
CC	send an email to license@isdb-sib.ch).
CC	EMBL; AL109732; GARS2005.1; -
DR	InterPro; IPR000370; Amylrolase
DR	InterPro; IPR000370; PTC_Amylrolase
DR	Pfam; PF00561; amydr0210; PTC_Amylrolase
DR	PRINTS; PR00793; PROAMNOPTASE
DR	Hydrolase; Aminopeptidase


```
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUGCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN; MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- DOMAIN: ONE OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright . It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or seed at email to license@isb.sdb.ch (see http://www.isb.sdb.ch/license/).
CC -----
CC DR EMBL; X65194; CAA6311.1; -.
CC DR PIR; S30259; S30259.
CC DR InterPro; IPR001067; MLC_carrier.
CC DR ProSITE; PS00215; MITOCH_CARRIER.
CC DR PRINTS; PR00927; ACPNRCCLASSE.
CC DR PRINTS; PR00927; ACPNRCCLASSE.
CC DR POSITE; PS00215; MITOCH_CARRIER; 2.
CC KW MITOCONDITION: Inner membrane. Repeat: Transmembrane; Transport.
CC FT TRANSMEM 12 92 1 (POTENTIAL).
CC FT TRANSMEM 74 29 2 (POTENTIAL).
CC FT TRANSMEM 116 133 3 (POTENTIAL).
CC FT TRANSMEM 116 133 3 (POTENTIAL).
CC FT TRANSMEM 219 237 4 (POTENTIAL).
CC FT TRANSMEM 273 281 6 (POTENTIAL).
CC FT TRANSMEM 308 AA: D47GFEQZB287A53F CRC64;
CC SEQUENCE 308 AA: 33528 WW:
CC -----
Cc Query Match 49.4%; Score 43; DB 1: Length 308;
Cc Best Local Similarity 53.3%; Pred. No. 32;
Cc Matches 8; Conservative 23; Mismatches 5; Indels 0; Caps 0;
Cc QY 2 SKLYVFLIWMGTWG 16
Cc Db 212 NFWLAFLGLMGITIG 226
Cc -----
Cc RESULT 9
Cc RESIDU: TOPPA
Cc ID OPSP TOOPA STANDARD: PROT: 447 AA.
Cc AC P13156;
Cc DT 01-JUL-1993 (Rel. 26, Created)
Cc DT 01-JUL-1993 (Rel. 26, Last sequence update)
Cc DT 15-JUN-1999 (Rel. 38, Last annotation update)
Cc GN RHODOPSIN.
Cc OS Rhodospirillum rubrum (Japanese flying squid).
Cc OR Rhodospirillum rubrum (Mollusca: Cephalopoda: Coleoidae; Teuthoidea;
Cc Oegopsida; Ammonostraphidae; Rhodactidea.
Cc NCBI_Taxid=6637;
Cc [1]
Cc RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-16; 128-139 AND 302-312.
Cc RA TISSUE: Retina;
Cc RL MEDLINE=93154520; PubMed=8428633;
Cc RA Hata-Mashimura T., Kondo M., Nishimura M., Hara R., Hara T.;
Cc RL Rhodopsin: The amino acid sequence of cDNA for rhodopsin of the squid
Cc RL Todarodes pacificus1.
Cc FEBS Lett. 317:5-11(1993).
Cc CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
Cc CC MEDiate VISION. THEY CONSIST OF AN APOLIPOPROTEIN, OPSIN, COVALENTLY
Cc CC LINKED TO CIS-RETINAL. THIS RECEPTOR IS COUPLED TO THE ACTIVATION
Cc CC OF PHOSPHOLIPASE C.
Cc CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
Cc CC -1- BE PHOSPHORYLATED.
Cc CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
Cc CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
Cc CC -----
Cc THIS SWISS-PROT entry is copyright . It is produced through a collaboration
```

CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed, usage by third parties for commercial
CC	or send an email to license@ebi.ac.uk or see http://www.ebi.ac.uk/infocentre/
CC	-- or send an email to license@sdb.sdb.ch
DR	EMBL; X7498; CAA4906.1 ;
DR	PIR; S29483; S29483.
DR	GCRB; GCR_0576; -
DR	InterPro; IPR000176; GPCR_Rhodopsin.
DR	InterPro; IPR001160; Opsin.
DR	InterPro; IPR001160; Opsin.
DR	Pfam; PF00001; 7tm.1.1-1.
DR	Pfam; PF02162; Rhodopsin_C; 1.
DR	PRINTS; PR00237; GPCRRHODOPSIN.
DR	PRINTS; PR00238; OPSIN.
DR	PRINTS; PR00239; RHODOPNITL.
DR	PROSITE; PS02277; G_PROTEIN_RECEP_F1_1; 1.
DR	PROSITE; PS02277; G_PROTEIN_RECEP_F1_2; 1.
DR	PROSITE; PS00228; Opsin; 1; N_KCNP_FL_2; 1.
KM	Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
KW	Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
FT	INT_MET 0 0
FT	DOMAIN 1 32
FT	TRANSMEM 33 57
FT	TRANSMEM 58 69
FT	TRANSMEM 70 81
FT	DOMAIN 82 97
FT	TRANSMEM 98 110
FT	TRANSMEM 111 130
FT	DOMAIN 131 150
FT	TRANSMEM 151 174
FT	TRANSMEM 175 198
FT	DOMAIN 199 226
FT	TRANSMEM 227 240
FT	TRANSMEM 241 260
FT	TRANSMEM 261 285
FT	DOMAIN 286 293
FT	TRANSMEM 294 317
FT	DOMAIN 318 447
FT	CARDHD 107 185
FT	DISULFD 107 185
FT	BLINDING 304 304
FT	LIPID 305 332
FT	LIPID 333 352
FT	LIPID 353 370
FT	DOMAIN 370 381
FT	DOMAIN 381 447
SO	SEQUENCE 447 AA; 49704 MW; ACTGADTFFE15A0IC1 CRC64;
Query Match	49.48; Score 43; DB 1; Length 447;
Best Local Similarity	38.5%; Pctd No. 43;
Matches 5; Conservative	3; Mismatches
OY 3 MYVFFLLMSYVM 15	
DB 153 IMIFWMYSVM 165	
RESURF 10	
PROD_SALT	
PROD_SALT	
AC 030646; STANDARD; PRT; 304 AA.	
DT 15-JUL-1999 (Rel. 38, Created)	
DT 15-JUL-1999 (Rel. 38, Last sequence update)	
DT 20-AUG-2001 (Rel. 40, Last annotation update)	
DE PAO PROTEIN.	
GN PAO.	
OS Salmonella typhimurium.	
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
CC Salmonella	
NCBI_Taxid=602;	
[1]	
NR SEQUENCE FROM N.A.	

```
CC RC STRAIN-ATCC 14028S;
CC MEDLINE-98360520; PubMed-9712687;
CC Guin J.S., Beiden W.J., Miller S.L.;
CC "Identification of Piro-Piro activated genes within a duplicated
CC chromosome from Piroplasma sp. strain 1000."
CC MISCdb Pathogen 25:77-80.(1998)
CC CC -1- SUPRACELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE EMAA TRANSPORTER FAMILY.
CC
CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use or distribution. However, if you wish to use it in a publication, please
CC modified and this statement is acknowledged as used by you. If you wish to
CC utilize requires a license agreement (see http://www.isdb.ch/announce/
CC or send an email to license@isdb-sib.ch)."http://www.isdb.ch/announce/
CC or send an email to license@isdb-sib.ch)."
```

```
CC CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by for commercial purposes requires a license agreement (see http://www.isb-sib.ch/announcements or send an email to license@isb-sib.ch).
CC CC .....
CC CC PMID: 862045; CNA4978.1; ..
CC CC PIR: S24055; S24065 ..
CC CC InterPro: IPR002208; secy..
CC CC Pfam: PF03044; secy_1..
CC CC PROSITE: PS00795; SECY_1..
CC CC PROSITE: PS00756; SECY_2; 1..
CC CC KM Protein transport; Translocation; Transmembrane..
CC CC FT TRANSMEM 28 48 POTENTIAL..
CC CC FT TRANSMEM 13 13 POTENTIAL..
CC CC FT TRANSMEM 113 133 POTENTIAL..
CC CC FT TRANSMEM 136 156 POTENTIAL..
CC CC FT TRANSMEM 168 188 POTENTIAL..
CC CC FT TRANSMEM 205 225 POTENTIAL..
CC CC FT TRANSMEM 257 277 POTENTIAL..
CC CC FT TRANSMEM 291 311 POTENTIAL..
CC CC FT TRANSMEM 316 336 POTENTIAL..
CC CC FT TRANSMEM 336 356 POTENTIAL..
CC CC FT TRANSMEM 356 376 POTENTIAL..
CC CC SQ SEQUENCE 438 AA: 47015 MW: 483925DD44E946A CAC64;
CC CC .....
Cc Db 323 MAFVCLFGLPFW 335
Cc Oy 3 MAYELLINGRUP 15
Cc |::|::|::|::|::|::|
Cc Best Local Similarity 48.3%; Score 42; DB 1; Length 438;
Cc Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Cc .....
Cc RESULT 12
Cc OPSP_OCTDO STANDARD: PERT: 455 AA.
Cc AD OCTDO OCTDO
Cc RA P03241
Cc RX 01-MAR-1989 (Rel. 10, Created)
Cc DT 01-MAR-1989 (Rel. 10, Last sequence update)
Cc DT 15-JUL-1999 (Rel. 38, Last annotation update)
Cc DE RHODOPSIN.
Cc GN RHO.
Cc OS Octopus doctlein (giant octopus).
Cc EC Ectopoda; Delezoza; Molusca; Cephalopoda; Coleutheta; Octopoda;
Cc NCBI Taxid=6644;
Cc OX NCBI_Taxid=6644;
Cc RN [1]
Cc RP SEQUENCE FROM N.A.
Cc MEDLINE=88211878; PubMed=3366250;
Cc RA Ovchinnikov Y.Y., Abdualeev N.G., Zolotarev A.S., Artemov I.D.,
Cc Bessipov I.A., Bergshoeff A.E., Tsuda M.;
Cc FEBS Lett. 232:69-72(1988).
Cc RL [2]
Cc RP REVISION TO 399.
Cc RA Abdualeev N.G.:
Cc Submitted (OCT-1988) to the EMBL/GenBank/DDBJ databases.
Cc RL [1]
Cc -! FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
Cc MODULATE VISION. THEY CONSIST OF AN APPROPRIATE, OSTIN, COVALENTLY
Cc -! SURCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
Cc -! TISSUE SPECIFICITY: MOD SHARPED PHOTORECEPTOR CELLS WHICH MEDIATES
Cc VISION IN DIM LIGHT.
Cc -! PM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
Cc BE PHOSPHORYLATED.
Cc -! SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
Cc -! OSTIN SUBFAMILY.
Cc .....
Cc This Swiss-Prot entry is copyright © It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation at
Cc the European Bioinformatics Institute. There are no restrictions on how
Cc it can be used, provided you acknowledge the source. You may not use it for
Cc use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: J01077, CAN50644.1; -.
 DR GENE: 00110, COG5.
 DR GCRDB: 007, 01007.
 DR InterPro: IPR000226; GPCR_Rhodopsin.
 DR InterPro: IPR000226; GPCR_Rhodopsin.
 DR InterPro: IPR001760; Opsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF02162; Rhodopsin_G; 1.
 DR PRINTS: PR00238; GPCR_RHODOPSIN.
 DR PRINTS: PR00239; RHODOPSIN.
 DR PROSITE: PS00237; G-PROTEIN-RECEPTOR-FL-1; 1.
 DR PROSITE: PS00262; G-PROTEIN-RECEPTOR-FL-2; 1.
 DR PROSITE: PS00283; OPSIN; 1.
 KM Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
 KM Photoreceptor; Retinal protein; Palmitate; G-protein coupled receptor;
 FT DOMAIN; 1 34
 FT TRANSMEM 35 59
 FT DOMAIN 60 71
 FT TRANSMEM 72 98
 FT DOMAIN 99 112
 FT TRANSMEM 113 132
 FT DOMAIN 133 152
 FT TRANSMEM 153 172
 FT DOMAIN 173 200
 FT TRANSMEM 201 228
 FT DOMAIN 229 262
 FT TRANSMEM 263 286
 FT DOMAIN 287 294
 FT TRANSMEM 295 319
 FT DOMAIN 320 455
 FT TRANSMEM 456 511
 FT CARBOHYD 1 5
 FT TRANSMEM 6 15
 FT BINDING 109 187
 FT LIPID 306 306
 FT LIPID 337 337
 FT LIPID 338 338
 FT LIPID 370 380
 FT DOMAIN 381 455
 FT TRANSMEM 456 511
 FT REPEAT 400 404
 FT REPEAT 412 416
 FT REPEAT 417 421
 FT REPEAT 422 426
 FT REPEAT 427 431
 SQ SEQUENCE 455 AA; 50488 MW; AECFAEC30ED1346C CRC64;

Query Match 48.3%; Score 42; DB 1; Length 455;
 Best Local Similarity 23.1%; Pred. No. 60;
 Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 MLVFFLLMGVYV 15
 DB 155 LVLFVFMKMSV 167

RESULT 13
 XASL_ECO57 STANDARD; PRT; 511 AA.
 AC P50229; 20-AUG-2001 (Rel. 40, Created)
 AC P50229; 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE AMINO ACID ANTI-PORTER (EXTREME ACID SENSITIVITY PROTEIN).
 OS XASL OR ACNS OR GADC OR 22216 OR ECS0297.
 GN Escherichia coli O157:H7.

CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=83334;
 RN 11
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21156231; PubMed=11258796;
 RX MEDLINE=21074935; PubMed=11205591;
 RA Petre N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick N.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobleck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamous K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R., *et al.* (2001).
 RA Nature 409:529-533(2001).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makiho K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Imai T., Nakano E., Hasegawa K., Kurita O., Tanaka M., Ode T.,
 RA Kihara S., Shibata T., Hattori M., Shinohara H.; *et al.* (2001).
 RA *Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.*
 RL DNA Res. 8:11-22(2001).
 CC -I- FUNCTION: PROBABLE AMINO ACID TRANSPORTER. INVOLVED IN EXTREME
 CC -I- ACID RESISTANCE (BY SIMILARITY).
 CC -I- PROTEIN LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -I- SIMILARITY). BELONGS TO THE XASL FAMILY OF TRANSPORTERS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AE005356; AAC56276.1; -.
 DR EMBL: AP002557; BAB35520.1; -.
 KM Transmembrane; Inner membrane; Transport; Amino-acid transport.
 FT TRANSMEM 15 35
 FT TRANSMEM 36 112
 FT TRANSMEM 113 148
 FT TRANSMEM 149 178
 FT TRANSMEM 201 221
 FT TRANSMEM 240 260
 FT TRANSMEM 292 312
 FT TRANSMEM 336 356
 FT TRANSMEM 357 433
 FT TRANSMEM 434 456
 FT TRANSMEM 457 466
 SQ SEQUENCE 511 AA; 55102 MW; B5DF7BD2529CE33 CRC64;

Query Match 48.3%; Score 42; DB 1; Length 511;
 Best Local Similarity 66.7%; Pred. No. 65;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LVFFLLMGVYV 15
 DB 42 LVFFLLMGVYV 53

RESULT 14
 XASL_ECO11 STANDARD; PRT; 511 AA.
 AC P39183; 01-NOV-1997 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

RESULT 16
 DALS, YEAST STANDARD: 543 AA.
 AC P15365.
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, last sequence update)
 DT 01-OCT-1996 (Rel. 34, last annotation update)
 DE ALLANTOATE PERMEASE.
 GN DALS OR UREP1 OR YJR152M OR J2230.
 OS Saccharomyces cerevisiae (Baker's yeast)
 CC Eukaryote; Eukaryota; Eukaryota; Saccharomycetes;
 CC Saccharomycetes; Saccharomycetes; Saccharomycetes;
 CC Saccharomycetes; Saccharomycetes; Saccharomycetes;
 CC NCBI_TaxID=4932.
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80086880; PubMed=3275614;
 RA Rai R., Genbaitte F.S., Cooper T.G.;
 RA Structure and transcription of the allantate permease gene (DALS)
 RA from *Saccharomyces cerevisiae*.
 RA J. Bacteriol. 170:266-271(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Scarce T.;
 RA Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: COMPONENT OF THE ALLANTOATE TRANSPORT SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE ALLANTOATE PERMEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR PMBL: M24098; AAA34555.1;
 DR EMBL: Z49652; CAB9685.1;
 DR PIR: A28671; A28671.
 DR SGP: S0003913; DALS.
 RN Transport: Transmembrane.
 KM TRANSMEM 81
 FT TRANSMEM 131
 FT TRANSMEM 145
 FT TRANSMEM 155
 FT TRANSMEM 179
 FT TRANSMEM 200
 FT TRANSMEM 214
 FT TRANSMEM 237
 FT TRANSMEM 249
 FT TRANSMEM 269
 FT TRANSMEM 318
 FT TRANSMEM 342
 FT TRANSMEM 353
 FT TRANSMEM 377
 FT TRANSMEM 418
 FT TRANSMEM 435
 FT TRANSMEM 454
 FT TRANSMEM 483
 FT TRANSMEM 504
 FT CARBOHYD 10
 FT CARBOHYD 19
 FT CARBOHYD 437
 FT CARBOHYD 441
 FT CARBOHYD 443
 FT SEQUENCE 543 AA; 60850 MW; 9870CASFALDDB88 CRGA;
 Query Match 48.3%; Score 42; DB 1; Length 543;
 Best Local Similarity 54.5%; Pctd. NB: 68;
 Matches 6; Consecutive 4; Mismatches 1; Indels 0; Gaps 0;
 Ov 3 MUYVYVIMCW 13
 Db 154 MUYVYVIMCW 164
 RESULT 17
 NTDO_CAEEL

ID NTDO_CAEEL STANDARD: 615 AA.
 AC 003614; 09XTKO.
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 20-AUG-2001 (Rel. 40, last sequence update)
 DT 01-APR-2001 (Rel. 40, last annotation update)
 DE SODIUM-DEPENDENT DOPAMINE TRANSPORTER (DA TRANSPORTER) (DAT).
 GN T32C5.5.
 OS Caenorhabditis elegans.
 CC Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Metazoa; Nematoda; Chromadorea; Rhabditidae;
 CC Rhabditidae; Metazoa; Nematoda; Chromadorea; Rhabditidae;
 CC NCBI_TaxID=6239.
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98440631; PubMed=9755501;
 RA Jayanthi L.D., Appasundaram S., Malone M.D., Ward E., Miller D.M.,
 RA Eppler M., Blakely R.D.;
 RA *The Caenorhabditis elegans gene t32c5.5 encodes an antidepressant and
 RA cocaine-sensitive dopamine transporter*;
 RA Mol. Pharmacol. 54:601-609(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Scarce T.;
 RA Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: TERMINATES THE ACTION OF DOPAMINE BY ITS HIGH
 CC AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE SODIUM-NEUROTRANSMITTER SYMPORTER
 CC FAMILY (SNF).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: AF079899; AAC3661.1;
 DR EMBL: Z19158; CAU7575.2;
 DR EMBL: Z19156; CAU7564.2;
 DR EMBL: Z19158; CAU7564.2; JOINED.
 DR PIR: S28306; S28306.
 DR WormPep: T32C5.5; CE25124.
 DR InterPro: IPR000175; Na_neurotran_symport.
 DR Pfam: PF00209; SNF; 1; Na_neurotran_symport.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1;
 DR PROSITE: PS00754; NA_NEUROTRAN_SYM_2; 1;
 DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1;
 KW Neurotransmitter Transport; Transport; Transmembrane; Glycoprotein;
 KW DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).

FT	TRANSMEM	74	96	1	(POTENTIAL).
FT	TRANSMEM	74	94	2	(POTENTIAL).
FT	TRANSMEM	125	145	3	(POTENTIAL).
FT	DONUTIN	246	282	4	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	246	252	5	(POTENTIAL).
FT	TRANSMEM	258	278	6	(POTENTIAL).
FT	TRANSMEM	344	364	7	(POTENTIAL).
FT	TRANSMEM	371	391	8	(POTENTIAL).
FT	TRANSMEM	393	413	9	(POTENTIAL).
FT	TRANSMEM	440	460	10	(POTENTIAL).
FT	TRANSMEM	462	482	11	(POTENTIAL).
FT	TRANSMEM	513	535	12	(POTENTIAL).
FT	TRANSMEM	535	555	13	(POTENTIAL).
FT	DOMAIN	571	615	14	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	162	182	15	N-LINKED (GLCNAC-...) (POTENTIAL).
FT	CARBOHYD	187	187	16	N-LINKED (GLCNAC-...) (POTENTIAL).
SQ	SEQUENCE	615 AA:	69265 MW:	DDA287EB3JC8049 CRC64:	

Query Match	48 3%	Score 42.	DB 1:	Length 615;
Best Local Similarity	35 7%	Pred No. 75		
Matches 5:	Conservative 7:	Mismatches 2:	Indels 0:	Gaps 0:

Oy	3	MLVFLILLCMCVNG 16	:::::
Db	463	LIMELMILLITVTW 476	

RESULT 18			
ID	KXIJ_SHIFL	STANDARD:	PRT: 241 AA.
AC	006081:		
Dt	01-JUN-1994 (Rel. 29, Created)		
Dt	01-JUN-1994 (Rel. 29, Last sequence update)		
Dt	01-FEB-1996 (Rel. 33, Last annotation update)		
Dl	PROTEIN KXIJ_PDBSUMSK.		
GN	KXIJ		
OS	Shigella flexneri.		
OC	Plasmid 210 Kb Invasin PWR100.		
BC	Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae;		
NCBI_TaxId=623:	Shigella.		
OR	NCBI_TaxId=623:		
OX	SEQUENCE FROM N.A.		
RX	STRAIN 900T.		
RC	MEDLINE:93077448; PubMed:1332940;		
RA	Alelou A., Sansonetti P.J., Parsot C.;		
RT	"KXIJ," a lipoprotein involved in secretion of Shigella ipa invasins,		
RT	is homologous to YscI, a secretion factor of the Yersinia yop		
RT	proteins". J. Bacteriol. 174:7651-7661(1992).		
SD	- SEQUENCE INVOLVED IN THE SECRETION OF THE IPA ANTIGENS		
SD	- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID		
SD	- ANCHOR (PROBABLY).		
SD	- SIMILARITY: BELONGS TO THE YSCJ FAMILY OF LIPOPROTEINS.		
SD	- THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
SD	between the Swiss Institute of Bioinformatics and the EMBL Outstation		
SD	at the University of Cambridge. There are no restrictions on its		
SD	use by non-profit institutions. However, there are no restrictions on its		
SD	modified and this statement is not removed, usage by and for commercial		
SD	entities requires a license agreement (See http://www.isb-sib.ch/announcements/		
SD	or send an email to license@lsb-sib.ch).		
DR	EMBL: M9B390: GA26532.1 -		
DR	PIR: G45271: G45271.		
DR	InterPro: IPR003520: SECTION1F.		
DR	InterPro: IPR003520: SECTION1F.		
DR	Plan: PF01514: YscJ.PLF.1		
DR	Prosite: PS00013: PROKR.LIPOPROTEIN_1.		
KW	Virulence; Plasmid; Transprot; Protein transport; Outer membrane;		
KW	Signal; Lipoprotein.		
ST	SIGNAL 1 17 POTENTIAL.		

FT	CHAIN	18	241	LIPROTEIN MXJL
FT	LIPID	18	18	N-ACYL DIGLYCERIDE (POTENTIAL)
SO	SEQUENCE	241 AA:	27509 MW; 302062033A30898 CRC64;	
<hr/>				
Query Match				
	Best Local Similarity	47.1%	Score 41:	DB 1:
	Matches 8:	Conservative	3:	Mismatches
OY	1 MSMLVFLLMGCVWG	16		Length 241:
	:::: :: ::	1		Prod. No. 49:
D6	219 MAYLVITLLWAFRTG	234		
<hr/>				
Query Match				
	Best Local Similarity	47.1%	Score 41:	DB 1:
	Matches 8:	Conservative	3:	Mismatches
OY	1 MSMLVFLLMGCVWG	16		Length 241:
	:::: :: ::	1		Prod. No. 49:
D6	219 MAYLVITLLWAFRTG	234		
<hr/>				
RESULT 19				
ID	MXJL_SHIISO	STANDARD:	PRT:	241 AA.
IC	055288:			
OC	NOV-1997 (rel. 35, created)			
DT	01-NOV-1997 (rel. 35, last sequence update)			
D7	01-NOV-1997 (rel. 35, annotation update)			
DE	LIPROTEIN MXJL PRECURSOR.			
CN	MXJL			
OS	Shigella sonnei.			
CC	Plasmid.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
CC	Shigella			
RX	NCBI_TaxID=624:			
RP	SEQUENCE FROM N.A.			
RA	SFRAIM-HM383:			
RA	AraKawa E., Kato J.I., Ito K.I., Watanabe H.;			
RL	Submitted (May-1995) to the EMBL/GenBank/DBJ databases.			
CC	-1 FUNCTION: INVOLVED IN THE SECRETION OF THE IPA ANTIGENS.			
CC	-1 SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID			
CC	ANCHOR (PROBABLE).			
CC	-1 SIMILARITY: BELONGS TO THE YSCJ FAMILY OF LIPOPROTEINS.			
CC	THIS SMS-S-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTRIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/)			
CC	(Contact:lsb-sib.ch)			
DR	EMBL: D50601: BAA09148.1: -			
DR	InterPro: IP8003283; SecEIOIMPX.			
DK	InterPro: IPR002920: YscJ_P1.F1.			
DK	Pfam: PF01514: YscJ_P1.F1.			
DR	PROSITE: PS00013: PROKR_LIPOPROTEIN; 1.			
KR	Intelligence: Plasmid Transport; Protein transport; Outer membrane;			
SM	Signal: Lipoprotein 17			
FT	CHAIN	18	241	POTENTIAL.
FT	LIPID	18	18	LIPROTEIN MXJL.
SO	SEQUENCE	241 AA:	27523 MW; 1FB3062033A31B20 CRC64;	
<hr/>				
Query Match				
	Best Local Similarity	50.0%	Score 41:	DB 1:
	Matches 8:	Conservative	3:	Mismatches
OY	1 MSMLVFLLMGCVWG	16		Length 241:
	:::: :: ::	1		Prod. No. 49:
D6	219 MAYLVITLLWAFRTG	234		
<hr/>				
RESULT 20				
ID	YJIE_ECOLI	STANDARD:	PRT:	301 AA.
IC	P32667:			
DT	01-OCT-1993 (rel. 27, created)			


```

DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
KW Hormone; Signal; Glycoprotein.
FT SIGNAL 23 42 BY SIMILARITY.
FT DISULFID 30 78 BY SIMILARITY.
FT DISULFID 44 93 BY SIMILARITY.
FT DISULFID 45 101 BY SIMILARITY.
FT DISULFID 55 139 BY SIMILARITY.
FT DISULFID 59 111 BY SIMILARITY.
FT DISULFID 114 121 BY SIMILARITY.
FT DISULFID 135 141 BY SIMILARITY.
SO SEQUENCE 141 AA; 15060 MW; 66CF803859C8EEDB CRC64;
Query Match 46.0%; Score 40; DB 1; Length 141;
Beat Local Similarity 43.8%; Pred. No. 44;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
Oy 1 MSMAPELLICCOLI 16
Db 7 LTVALILLILLRGGSK 22
RESULT 26
YFHB_ECOLI STANDARD: PRT; 211 AA.
ID YFHB_ECOLI
DR 01-APR-1993 (Ref. 25, Created)
DT 30-MAY-2000 (Ref. 39, Last sequence update)
DT 20-MAY-2001 (Ref. 40, Last annotation update)
DE HYPOHERETICAL 24.4 KDA PROTEIN IN PUBL-DEU INTERGENIC REGION.
OS YFHB OR B2560.
GN Escherichia coli.
GC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
GX NCBI_TaxID=562;
OX NCBI_TaxID=562;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=NM137.
RX MEDLINE:92292954; PubMed:1602968;
RX Poulsen L.K., Larsen N.M., Molin S., Andersson P.;
RX Bacterial strains resistant to the cell-
RX killing function encoded by the yef gene family.
RX Mol. Microbiol. 6:893-905(1992).
RL 12
RC SEQUENCE FROM N.A.
RC STRAIN=K12.
NA Nishimoto H., Saito N.;
NA Submitted (Sep-1995) to the EMBL/GenBank/ODDB databases.
RP 13
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MC155;
RX MEDLINE:97426617; PubMed:9278503;
RX Blatherer F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RX Riley M., Colado-Vides J., Clamen J.D., Rode C.K., Mayhew G.F.,
RX Gregor J., Davis N.W., Kirtpatrick H.A., Goeden M.A., Rose D.J.,
RX Mau B., Shao Y.;
RX The complete genome sequence of Escherichia coli K-12.
RL Science 277:1453-1474(1997)
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announc/.)
CC
CC EMBL: X72336; CA51063.1; ALT-INT.
CC EMBL: D64044; BA10510.1; ALT-INT.
CC EMBL: U09841; AA079822.1; ALT-INT.
CC EMBL: AE000342; AAC75613.1; ALT-INT.

```

```

DR      PIR: S20973; S20973.  
DR      Ecocore: Ecoli371; yfNB.  
KW      Hypothetical protein; Complete proteome.  
SQ      SEQUENCE   211 AA; 24438 MW;  90B93F5E34AAAD CRC64;  
  
Query Match          46.0%; Score 40.; DB 1; Length 211;  
Best local similarity 77.8%; Pred. No. 60;  
Matches    7; Conservative 1; Mismatches     1; Indels    0; Gaps    0;  
  
Db         65 ULMGNGV 73  
            11111..11  
            |||||...  
OY          8 LLMGNGV 16  
             :||:::  
DB          65 ULMGNGV 73  
              |:::  
              |||||..  
  
RESULT 27  
REFAL_KLEPN STANDARD; PRT; 259 AA.  
ID        REFAL_KLEPN  
AC        Q18475-1697 (Ref. 35. Created)  
FT        CDS       1-NOV-1997 (last update) [incomplete sequence]  
DT        01-NOV-1997 (Ref. 35. Last annotation update)  
DE        O-ANTIGEN EXPORT SYSTEM PERMEASE PROTEIN PEPA.  
CN        PEPA.  
OS        Klebsiella pneumoniae.  
OC        Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC        Klebsiella.  
CB        NCBI-TaxId=573;  
RP        SOURCE: FROM N.A.  
RX        STRAIN: 01-R20 / 889/50;  
RC        MEDLINE: 95191401; PubMed: 7533882;  
RA        Bromer D., Clarke B.R., Whitfield C.;  
RT        "Identification of an ATP-binding cassette transport system required  
for translocation of lipopolysaccharide O-antigen side-chains across  
the cytoplasmic membrane of Klebsiella pneumoniae serotype O1.";  
RM        Mol Microbiol 19:195-205(1997).  
CC        -1 FUNCTION: MAY FORM AN ATP-DRIVEN O-ANTIGEN EXPORT APPARATUS, IN  
ASSOCIATION WITH REPB.  
CC        -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
(POTENTIAL).  
CC        -1 SIMILARITY: BELONGS TO THE ABC-2 SUSPAMILY OF INTEGRAL MEMBRANE  
PROTEINS.....  
CC        THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation at  
The European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/donnce/  
or send an email to licencessib@sib.ch).  
DB        EMBL: L13775; AAC98411.1;  
DR        InterPro: IPR000412; ABC2_transport.  
RW        Pfam: PF01061; ABC2_membrane_1.  
KW        POSITIVE: PS00890; ABC2_MEMBRANE; 1.  
KW        Polysaccharide transport; Transport; Transmembrane; Inner membrane.  
FT        TRANSMEM 33      53      POTENTIAL.  
FT        TRANSMEM 59      113     POTENTIAL.  
FT        TRANSMEM 113     162     POTENTIAL.  
FT        TRANSHEM 142     162     POTENTIAL.  
FT        TRANSEM  176     196     POTENTIAL.  
FT        TRANSEM  228     248     POTENTIAL.  
SQ      SEQUENCE   259 AA; 30036 MW;  074811E59BA5C493 CRC64;  
  
Query Match          46.0%; Score 40.; DB 1; Length 259;  
Best local similarity 77.8%; Pred. No. 71;  
Matches    8; Conservative 5; Mismatches     1; Indels    0; Gaps    1;  
  
Db         3 MLVEFLMGCV-----NC 16  
           ::||||::|::|::|  
OY          3 MLVEFLMGCV-----NC 16  
           ::||||::|::|::|  
DB          124 YIVVELPYGVQMPSLMWGC 143
```


FT DOMAIN 254 261 EXTRACELLULAR (POTENTIAL).
 FT TRANSNM 262 286 7 (POTENTIAL).
 FT DOMAIN 287 >304 CYTOSOLIC (POTENTIAL).
 FT DISULFO 287 164 BT SIMILARITY.
 FT TRANSNM 288 293 BT SIMILARITY.
 FT LIPID 299 299 PALMITATE (BY SIMILARITY).
 FT LIPID 300 300 PALMITATE (BY SIMILARITY).
 FT NON_TER 304 304
 SQ SEQUENCE 304 AA: B1E9A6977136F46 CRC64:
 Query Match 46.0% Score 40; DB 1; Length 304;
 Best Local Similarity 43.2% Pred. No. 80;
 Matches 7; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
 Oy 3 MIVFILLM--GVYV 15
 1: ||| |||
 Db 234 MVIAPLDCMPCYAGVAV 250
 RESULT 30
 ADT_KLULA STANDARD: PRT: 305 AA.
 AC P49382:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE ADP-FLY-138-PROTEIN (ADP/ATP TRANSLOCASE) (ADENINE NUCLEOTIDE
 TRANSLOCATOR) (ANT).
 CN AAC.
 OS Kluyveromyces fragilis (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 RN 111
 RL NCBI_TaxID:28985;
 RP SEQUENCE FROM N.A.
 ME MEDLINE:95254654; PubMed:7736606;
 RA Viola A.M., Galeotti C.L., Goffelli P., Picarelli A., Ferrero I.;
 * A Kluyveromyces fragilis gene homologue to AAC2 complements the
 Saccharomycetes cerevisiae op1 mutation.*
 CC Cyt. Genet. 27:223-223(1993).
 CC -1- INHIBITS THE EXCHANGE OF ADP AND ATP ACROSS THE
 MITOCHONDRIAL INNER MEMBRANE.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 INNER MEMBRANE.
 CC -1- DOMAIN: COMPOSED OF THREE MITOCHONDRIAL CARRIER FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: L13797; AAC1655.1;
 DR InterPro: IPR002067; MitocCarrier.
 DR InterPro: IPR001993; MitocCarrier.
 DR Pfam: PF00153; mitocCarrier_3.
 DR PRINTS: PR00926; MITOCARRIER.
 DR PRINTS: PR00925; ADPTRNSLCASE.
 KM Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSNM 76 84 2 (POTENTIAL).
 FT TRANSNM 84 94 2 (POTENTIAL).
 FT TRANSNM 118 135 3 (POTENTIAL).
 FT TRANSNM 179 198 4 (POTENTIAL).
 FT TRANSNM 218 235 5 (POTENTIAL).
 FT TRANSNM 274 292 6 (POTENTIAL).
 SQ SEQUENCE 305 AA: 7F9596CC184B1A1 CRC64:

Query Match 46.0% Score 40; DB 1; Length 305;
 Best Local Similarity 60.0% Pred. No. 81;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Oy 2 SMVVFLLMCTYMG 16
 1: ||| ||| |||
 Db 213 SFLASFLGMAVYTG 227
 RESULT 31
 BTR_SALT STANDARD: PRT: 319 AA.
 AC Q05600:
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CRIB PROTEIN.
 CN GRIH.
 OS Salmonella typhimurium.
 CC Gram-negative proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CX NCBI_TaxID:602;
 RN 111
 RP SEQUENCE FROM N.A.
 CC STRAIN U72.
 RX MEDLINE:93272696; PubMed:9501034;
 RA Church G.M., Lawrence J.G., Rudenfeld M., Kieffer-Higgins S.,
 *Characterization of the cobalamin (vitamin B12) biosynthetic genes
 of Salmonella typhimurium*;
 RL J. Bacteriol. 175:3303-3316(1993).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF COBAYRIC ACID TO
 COBALAMINE. ADDITION OF AMINOPROPANOL ON THE F CARBOXYLIC
 GROUP.
 CC -1- PATHWAY: COBALAMIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE COB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: L12006; AAA7255.1;
 DR Strydom; S010053; GRIH.
 KM Outer membrane; Outer membrane; Transmembrane.
 FT TRANSNM 56 76 7 (POTENTIAL).
 FT TRANSNM 82 102 7 (POTENTIAL).
 FT TRANSNM 153 173 7 (POTENTIAL).
 FT TRANSNM 296 316 7 (POTENTIAL).
 SQ SEQUENCE 319 AA: 540DF95E0796214 CRC64:

Query Match 46.0% Score 40; DB 1; Length 319;
 Best Local Similarity 53.8% Pred. No. 83;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 4 LVVFLLMCTYMG 16
 1: || ||| |||
 Db 59 VVVVGATGCVAG 71
 RESULT 32
 YHHT_ECOLI STANDARD: PRT: 349 AA.
 AC P37622; P76700;
 DT 01-OCT-1994 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

RL      SubmitLink (JAN-1999) to the EMBL/Genbank/ODD databases.
CC      -1- FUNCTION: VISUAL ELEMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
CC      MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC      LINKED TO CIS-RETINAL.
CC      -1- FUNCTION: VISUAL ELEMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
CC      MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC      LINKED TO CIS-RETINAL.
CC      -1- TISSUE SPECIFICITY: ROD SHAPED PHOTORECEPTOR CELLS WHICH MEDIATES
CC      VISION IN DIM LIGHT.
CC      -1- PM: SOME OR ALL OF THE CARBOXY-TERMINAL SER OR THR RESIDUES MAY
CC      BE PHOSPHORYLATED.
CC      -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC      OPSIN SUBFAMILY.
CC
CC      THIS SWISS-PROT ENTRY IS COURTESY IT IS PRODUCED THROUGH A COLLABORATION
CC      BETWEEN THE Swiss Institute of Bioinformatics and the EMBL contribution
CC      to the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.lsb-sib.ch/announcements/
CC      or send an email to license@lsb-sib.ch).
CC
CC      EMBL: Y18669; CAJ73251.1
CC      InterPro: IPRO00276; GPCR_Rhodopsin.
CC      InterPro: IPRO01760; Opsin.
CC      Pfam: PF00001; 7tm.1.1.
CC      PRINTS: PR00237; GPCR_RHODOPSIN.
CC      PRINTS: PR00238; OPSIN.
CC      PRINTS: PR00239; RHODOPSIN.
CC      PROSITE: PS00227; G_PROTEIN_RECEP_F1_1; 1.
CC      PROSITE: PS00228; G_PROTEIN_RECEP_F1_2; 1.
CC      PROSITE: PS00228; G_PROTEIN_RECEP_F1_2; 1.
CC      KW      Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
CC      Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
CC
CC      FT      DOMAIN 1 36
CC      FT      TRANSMEM 37 61 1 (POTENTIAL).
CC      FT      TRANSMEM 62 73 2 (POTENTIAL).
CC      FT      TRANSMEM 74 96 3 (POTENTIAL).
CC      FT      TRANSMEM 97 113 4 (POTENTIAL).
CC      FT      TRANSMEM 114 133 5 (POTENTIAL).
CC      FT      TRANSMEM 134 152 6 (POTENTIAL).
CC      FT      TRANSMEM 153 176 7 (POTENTIAL).
CC      FT      DOMAIN 177 202 8 (POTENTIAL).
CC      FT      TRANSMEM 203 230 9 (POTENTIAL).
CC      FT      TRANSMEM 231 252 10 (POTENTIAL).
CC      FT      TRANSMEM 253 273 11 (POTENTIAL).
CC      FT      TRANSMEM 274 305 12 (POTENTIAL).
CC      FT      TRANSMEM 306 325 13 (POTENTIAL).
CC      FT      TRANSMEM 326 353 14 (POTENTIAL).
CC      FT      DOMAIN 354 379 15 (POTENTIAL).
CC      FT      CARBOHYD 15 15 15 (POTENTIAL).
CC      FT      CARBOHYD 200 200 200 (POTENTIAL).
CC      FT      BINDING 295 295 295 (POTENTIAL).
CC      FT      BINDING 326 326 326 (POTENTIAL).
CC      FT      DISULFID 110 187 187 (POTENTIAL).
CC      FT      DISULFID 110 187 187 (POTENTIAL).
CC      SEQUENCE 353 AA: 39505 MW: 8CB6FA847B3C32 CRC64:
CC
Query Match 46.08: Score 40; DB 1; Length 353;
Best Local Similarity 47.18: Pred. No. 50;
Matches 8; Conservative 2; Mismatches 3; Indels 4; Gaps 1.
Oy 3 MCVYVILLM---GVN 15
Db 257 NYVAVLLEKPPYAGVAV 273

```

GN RHIO. Lithognathus mormyrus (Striped seabream).
 CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 CC Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neocosteleostei:
 CC Clupeiformes: Clupeidae: Sardina: Sardina (Sardine).
 CC Striatidae: Striatidae: Sardina: Sardina (Sardine).
 CC NCBI_TaxID:505951.
 OK [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE: Retina.
 RA Archer S.N., Hwang J.J. opsin in *Mellieriana* coaxed fish. *
 RA Comparative biochemistry and physiology. Part B. 1996;115(1):1-10.
 RT F1 FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
 CC LINKED TO CIS-RETINAL.
 CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1 TISSUE SPECIFICITY: ROD SHAPED PHOTORECEPTOR CELLS WHICH MEDIANES
 CC VISION IN DIM LIGHT.
 CC -1 BR PHOSPHORYLATED OF THE CARBOXY-TERMINAL SER OR THR RESIDUES MAY
 CC -1 SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC ORIGIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation *
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use, but the EMBL logo may not be used as a trade mark by and for commercial
 CC entities requires a license agreement (see <http://www.embnet.org/announcements/>
 CC or send an email to license@ebi.scrib.ch).
 CC -----
 DR EMBL: Y18667; CAAY7249.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00010; Rhodopsin.
 DR Pfam: PF00010; Rhodopsin.
 DR PRINTS: PR00237; GPCR_RHO000SN.
 DR PRINTS: PR00238; OPSIN.
 DR PRINTS: PR00579; RHODOPSIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1;
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2;
 DR PROSITE: PS00288; OPSIN.
 DR PROSITE: PS00289; OPSIN.
 DR PROSITE: PS00290; OPSIN.
 DR PROSITE: PS00291; OPSIN.
 DR PROSITE: PS00292; OPSIN.
 DR PROSITE: PS00293; OPSIN.
 DR PROSITE: PS00294; OPSIN.
 DR PROSITE: PS00295; OPSIN.
 DR PROSITE: PS00296; OPSIN.
 DR PROSITE: PS00297; OPSIN.
 DR PROSITE: PS00298; OPSIN.
 DR PROSITE: PS00299; OPSIN.
 DR PROSITE: PS00300; OPSIN.
 DR PROSITE: PS00301; OPSIN.
 DR PROSITE: PS00302; OPSIN.
 DR PROSITE: PS00303; OPSIN.
 DR PROSITE: PS00304; OPSIN.
 DR PROSITE: PS00305; OPSIN.
 DR PROSITE: PS00306; OPSIN.
 DR PROSITE: PS00307; OPSIN.
 DR PROSITE: PS00308; OPSIN.
 DR PROSITE: PS00309; OPSIN.
 DR PROSITE: PS00310; OPSIN.
 DR PROSITE: PS00311; OPSIN.
 DR PROSITE: PS00312; OPSIN.
 DR PROSITE: PS00313; OPSIN.
 DR PROSITE: PS00314; OPSIN.
 DR PROSITE: PS00315; OPSIN.
 DR PROSITE: PS00316; OPSIN.
 DR PROSITE: PS00317; OPSIN.
 DR PROSITE: PS00318; OPSIN.
 DR PROSITE: PS00319; OPSIN.
 DR PROSITE: PS00320; OPSIN.
 DR PROSITE: PS00321; OPSIN.
 DR PROSITE: PS00322; OPSIN.
 DR PROSITE: PS00323; OPSIN.
 DR PROSITE: PS00324; OPSIN.
 DR PROSITE: PS00325; OPSIN.
 DR PROSITE: PS00326; OPSIN.
 DR PROSITE: PS00327; OPSIN.
 DR PROSITE: PS00328; OPSIN.
 DR PROSITE: PS00329; OPSIN.
 DR PROSITE: PS00330; OPSIN.
 DR PROSITE: PS00331; OPSIN.
 DR PROSITE: PS00332; OPSIN.
 DR PROSITE: PS00333; OPSIN.
 DR PROSITE: PS00334; OPSIN.
 DR PROSITE: PS00335; OPSIN.
 DR PROSITE: PS00336; OPSIN.
 DR PROSITE: PS00337; OPSIN.
 DR PROSITE: PS00338; OPSIN.
 DR PROSITE: PS00339; OPSIN.
 DR PROSITE: PS00340; OPSIN.
 DR PROSITE: PS00341; OPSIN.
 DR PROSITE: PS00342; OPSIN.
 DR PROSITE: PS00343; OPSIN.
 DR PROSITE: PS00344; OPSIN.
 DR PROSITE: PS00345; OPSIN.
 DR PROSITE: PS00346; OPSIN.
 DR PROSITE: PS00347; OPSIN.
 DR PROSITE: PS00348; OPSIN.
 DR PROSITE: PS00349; OPSIN.
 DR PROSITE: PS00350; OPSIN.
 DR PROSITE: PS00351; OPSIN.
 DR PROSITE: PS00352; OPSIN.
 DR PROSITE: PS00353; OPSIN.
 DR PROSITE: PS00354; OPSIN.
 DR PROSITE: PS00355; OPSIN.
 DR PROSITE: PS00356; OPSIN.
 DR PROSITE: PS00357; OPSIN.
 DR PROSITE: PS00358; OPSIN.
 DR PROSITE: PS00359; OPSIN.
 DR PROSITE: PS00360; OPSIN.
 DR PROSITE: PS00361; OPSIN.
 DR PROSITE: PS00362; OPSIN.
 DR PROSITE: PS00363; OPSIN.
 DR PROSITE: PS00364; OPSIN.
 DR PROSITE: PS00365; OPSIN.
 DR PROSITE: PS00366; OPSIN.
 DR PROSITE: PS00367; OPSIN.
 DR PROSITE: PS00368; OPSIN.
 DR PROSITE: PS00369; OPSIN.
 DR PROSITE: PS00370; OPSIN.
 DR PROSITE: PS00371; OPSIN.
 DR PROSITE: PS00372; OPSIN.
 DR PROSITE: PS00373; OPSIN.
 DR PROSITE: PS00374; OPSIN.
 DR PROSITE: PS00375; OPSIN.
 DR PROSITE: PS00376; OPSIN.
 DR PROSITE: PS00377; OPSIN.
 DR PROSITE: PS00378; OPSIN.
 DR PROSITE: PS00379; OPSIN.
 DR PROSITE: PS00380; OPSIN.
 DR PROSITE: PS00381; OPSIN.
 DR PROSITE: PS00382; OPSIN.
 DR PROSITE: PS00383; OPSIN.
 DR PROSITE: PS00384; OPSIN.
 DR PROSITE: PS00385; OPSIN.
 DR PROSITE: PS00386; OPSIN.
 DR PROSITE: PS00387; OPSIN.
 DR PROSITE: PS00388; OPSIN.
 DR PROSITE: PS00389; OPSIN.
 DR PROSITE: PS00390; OPSIN.
 DR PROSITE: PS00391; OPSIN.
 DR PROSITE: PS00392; OPSIN.
 DR PROSITE: PS00393; OPSIN.
 DR PROSITE: PS00394; OPSIN.
 DR PROSITE: PS00395; OPSIN.
 DR PROSITE: PS00396; OPSIN.
 DR PROSITE: PS00397; OPSIN.
 DR PROSITE: PS00398; OPSIN.
 DR PROSITE: PS00399; OPSIN.
 DR PROSITE: PS00400; OPSIN.
 DR PROSITE: PS00401; OPSIN.
 DR PROSITE: PS00402; OPSIN.
 DR PROSITE: PS00403; OPSIN.
 DR PROSITE: PS00404; OPSIN.
 DR PROSITE: PS00405; OPSIN.
 DR PROSITE: PS00406; OPSIN.
 DR PROSITE: PS00407; OPSIN.
 DR PROSITE: PS00408; OPSIN.
 DR PROSITE: PS00409; OPSIN.
 DR PROSITE: PS00410; OPSIN.
 DR PROSITE: PS00411; OPSIN.
 DR PROSITE: PS00412; OPSIN.
 DR PROSITE: PS00413; OPSIN.
 DR PROSITE: PS00414; OPSIN.
 DR PROSITE: PS00415; OPSIN.
 DR PROSITE: PS00416; OPSIN.
 DR PROSITE: PS00417; OPSIN.
 DR PROSITE: PS00418; OPSIN.
 DR PROSITE: PS00419; OPSIN.
 DR PROSITE: PS00420; OPSIN.
 DR PROSITE: PS00421; OPSIN.
 DR PROSITE: PS00422; OPSIN.
 DR PROSITE: PS00423; OPSIN.
 DR PROSITE: PS00424; OPSIN.
 DR PROSITE: PS00425; OPSIN.
 DR PROSITE: PS00426; OPSIN.
 DR PROSITE: PS00427; OPSIN.
 DR PROSITE: PS00428; OPSIN.
 DR PROSITE: PS00429; OPSIN.
 DR PROSITE: PS00430; OPSIN.
 DR PROSITE: PS00431; OPSIN.
 DR PROSITE: PS00432; OPSIN.
 DR PROSITE: PS00433; OPSIN.
 DR PROSITE: PS00434; OPSIN.
 DR PROSITE: PS00435; OPSIN.
 DR PROSITE: PS00436; OPSIN.
 DR PROSITE: PS00437; OPSIN.
 DR PROSITE: PS00438; OPSIN.
 DR PROSITE: PS00439; OPSIN.
 DR PROSITE: PS00440; OPSIN.
 DR PROSITE: PS00441; OPSIN.
 DR PROSITE: PS00442; OPSIN.
 DR PROSITE: PS00443; OPSIN.
 DR PROSITE: PS00444; OPSIN.
 DR PROSITE: PS00445; OPSIN.
 DR PROSITE: PS00446; OPSIN.
 DR PROSITE: PS00447; OPSIN.
 DR PROSITE: PS00448; OPSIN.
 DR PROSITE: PS00449; OPSIN.
 DR PROSITE: PS00450; OPSIN.
 DR PROSITE: PS00451; OPSIN.
 DR PROSITE: PS00452; OPSIN.
 DR PROSITE: PS00453; OPSIN.
 DR PROSITE: PS00454; OPSIN.
 DR PROSITE: PS00455; OPSIN.
 DR PROSITE: PS00456; OPSIN.
 DR PROSITE: PS00457; OPSIN.
 DR PROSITE: PS00458; OPSIN.
 DR PROSITE: PS00459; OPSIN.
 DR PROSITE: PS00460; OPSIN.
 DR PROSITE: PS00461; OPSIN.
 DR PROSITE: PS00462; OPSIN.
 DR PROSITE: PS00463; OPSIN.
 DR PROSITE: PS00464; OPSIN.
 DR PROSITE: PS00465; OPSIN.
 DR PROSITE: PS00466; OPSIN.
 DR PROSITE: PS0046

[illegible]

Query Match 46.0% Score 40; DB 1; Length 353;
Best Local Similarity 47.1% Pred. No. 90;
Matches 8: Conservative 2; Mismatches 3; Indels 4; Gaps 1;

OY 3 MCVFLLM---GYW 15
DB 257 MVAFLMCPVAGYAM 273

RESULT 36
OPSD,SOLSO STANDARD; PRT: 353 AA.

AC 09YGS-2001 (rel. 40, Created)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE RHODOPSIN.
GN RHO.
OS Solea solea (Common sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Clupeiformes; Acanthopterygii; Perciformes; Pleuronectiformes;
OC Soleidae; Solea.
NC1_TaxID=90069;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Archer S.N., Altano J.;
RA Comparative analysis of opsins in Mediterranean coastal fish.*
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. The sequence is provided as a reference for the scientific community
CC modified and this statement is not removed or changed by and for commercial
CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC
CC EMBL: Y18672; CA77254.1;
CC InterPro: IPR001266; GPCR_Rhodopsin.
CC Pfam: PF00001; 7tm1; Opnln.
DR PRINTS: PRO0237; GPCR_Rhodopsin.
DR PRINTS: PRO0238; OPSIN.
DR PRINTS: PRO0579; RHODOPSIN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
KN Phosphorylation: Lipidation: Palmitate: G-protein coupled receptor.
FT DOMAIN 1 36
FT TRANSMEM 37 61
FT DOMAIN 62 73
FT TRANSMEM 74 98
FT DOMAIN 99 113
EXTRACELLULAR.
EXTRACELLULAR.

FT TRANSMEM 114 133
FT DOMAIN 134 152
FT TRANSMEM 153 176
FT DOMAIN 177 202
FT TRANSMEM 203 230
FT TRANSMEM 231 256
FT TRANSMEM 257 284
FT TRANSMEM 285 309
FT DOMAIN 310 353
FT CARBOHD 15 2
FT CARBOHD 200 15
FT CARBOHD 200 200
FT BINDING 305 325
FT BINDING 325 325
FT DISULFD 110 187
SEQUENCE 353 AA; B23570A75EB20FA CRC64;

Query Match 46.0% Score 40; DB 1; Length 353;
Best Local Similarity 41.2% Pred. No. 90;
Matches 7: Conservative 3; Mismatches 3; Indels 4; Gaps 1;

OY 3 MCVFLLM---GYW 15
DB 257 MVAFLMCPVAGYAM 273

RESULT 37
OPSD,SOLSO STANDARD; PRT: 354 AA.

AC P35359;
DT 01-JUN-1994 (rel. 29, Created)
DT 01-JUN-1994 (rel. 29, Last annotation update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE ULTRAVIOLET-SENSITIVE OPSIN (ULTRAVIOLET CONE PHOTORECEPTOR PIGMENT)
GN ZRO2.1.
OS Brachydanio rerio (zebrafish) (Zabira danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
NC1_TaxID=7955;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Robinson J., Schmitt E.A., Harosi F.I., Reese R.J., Dowling J.E.;
RA "Zebrafish ultraviolet visual pigment: absorption spectrum, sequence,
RA and localization".
RA Proc. Natl. Acad. Sci. U.S.A. 90:6009-6012(1993).
RP
RP SEQUENCE FROM N.A.
RC MEDLINE=69244282; PubMed=6924413;
RA Robinson J., Schmitt E.A., Dowling J.E.;
RA "Temporal and spatial patterns of opsin gene expression in zebrafish
RA (Danio rerio)".
RA Vis. Neurosci. 13:895-906(1995).
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. The sequence is provided as a reference for the scientific community
CC modified and this statement is not removed or changed by and for commercial
CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC
CC EMBL: Z00210; Z00210.1;
CC InterPro: IPR001266; GPCR_Rhodopsin.
CC Pfam: PF00001; 7tm1; Opnln.
DR PRINTS: PRO0237; GPCR_Rhodopsin.
DR PRINTS: PRO0238; OPSIN.
DR PRINTS: PRO0579; RHODOPSIN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
KN Phosphorylation: Lipidation: Palmitate: G-protein coupled receptor.
FT DOMAIN 1 36
FT TRANSMEM 37 61
FT DOMAIN 62 73
FT TRANSMEM 74 98
FT DOMAIN 99 113
EXTRACELLULAR.
EXTRACELLULAR.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M69160: AAA25866.1: -
 DR PDB: 1JGQ: 18-MAY-99.
 DR Lyase: 3D-structure.
 SW SEQUENCE 451 AA: 49572 MW: 04C85639BC1GCEA CRC64:

Query Match 46.0%: Score 40: DB 1: Length 451:
 Best Local Similarity 83.3%: Pred. No. 116+02:
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

OY 11 MOVING 16
 338 MOWING 343

RESULT 40
 ID GUDH_BACSU STANDARD: PRT: 455 AA.

AC P42338: (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-NOV-1997 (Rel. 35, Last annotation update)
 DE PROBABLE GLUCONATE DEHYDROGENASE (EC 4.2.1.40) (GUDH) (GUDC).
 GN GUDH.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423:

RP SEQUENCE FROM N.A.

RC STRAIN=168:

RE MEDLINE=95210079: PubMed-7704254:

RT "Determination of a 21548 bp nucleotide sequence around the 24

degrees region of the Bacillus subtilis chromosome.";

RL Microbiology 141:269-279(1995).

CC -1- DEXY-D-GLUCARATE 5-INDOLE-3-PYRROLIDINE CARBOXYLIC ACID
 CC -1- CATALYTIC ACTIVITY: D-GLUCARATE + 5-DEHYDRO-4-DEOXY-D-GLUCARATE +

H(2)O.

CC -1- PATHWAY: FIRST STEP IN GLUCONATE CATABOLISM.

CC -1- SIMILARITY: BELONGS TO THE NADPHASE RACEMASE / MUCONATE

LACTONIZING ENZYME FAMILY. GUDC SUBFAMILY.

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: 299105: G92043.1: -

DR Subtilisin: B01161: gnd.

SW Lyase: Complete proteome.
 KM SEQUENCE 455 AA: 50782 MW: 3238486007598C2A CRC64:

Query Match 46.0%: Score 40: DB 1: Length 455:
 Best Local Similarity 83.3%: Pred. No. 116+02:
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

OY 11 MOVING 16
 342 MOWING 347

RESULT 41
 ID FHUB_ECOLI STANDARD: PRT: 660 AA.

DT 01-APR-1988 (Rel. 07, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE FERRICHRON TRANSPORT SYSTEM PERMEASE PROTEIN FHUB (FERRICHRON UPTAKE
 DE PROTEIN FHUB).
 OS FHUB OR B0153.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia coli.
 OX NCBI_TaxID=562:

RP SEQUENCE FROM N.A.
 RE MEDLINE=87014116: PubMed-3020380:
 RT "Iron hydroxamate transport of Escherichia coli: nucleotide sequence
 RT of the protein and identification of the protein.";
 RL Mol. Gen. Genet. 204:435-442(1996).

RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110:
 RE MEDLINE=94261430: PubMed-8202364:
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RA Systematic sequencing of the Escherichia coli genome: analysis of
 RA the complete genome sequence of Escherichia coli K-12.
 RL Nucleic Acids Res. 22:1637-1639(1994).

RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655:
 RE MEDLINE=97426617: PubMed-9278503:
 RA Blatter F.R., Plunkett G. T.I., Bloch C.A., Berna N.T., Burland V.,
 RA Georg J., Davis N.W., Klaparek H.A., Gorden M.A., Rose D.J.,
 RA Men B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).

RP SEQUENCE FROM N.A.
 RA Schramm S., Duncan M., Allen E., Areujo R., Aparicio A., Chung E.,
 RA Lacharme F., Federespele N., Yamana R., Kallman R., Komp D.O.,
 RA Lashari D., Lee H., Lin D., Hamachi A., Oetler F., Roberts D.,
 RA Davis R.W.;
 RL Submitted (Sep-1996) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 1-23 FROM N.A.
 RE MEDLINE=88038363: PubMed-2823072:
 RA Burkhardt R., Braun V.; the fnc and fnd genes involved in iron
 RT (iii) hydroxamate transport: domains in Phoc homologous to
 RT ATP-binding proteins.";
 RL Mol. Gen. Genet. 209:49-55(1987).

CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS INVOLVED IN THE HIGH-
 CC AFFINITY TRANSPORT OF IRON(III)-FERRICHRON INTO THE E. COLI
 CC CELL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE FERRICHRON TRANSPORT SYSTEM. BELONGS TO THE FERRICHRON
 CC SUBFAMILY.

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X04319: CAA27852.1: -

DR EMBL: X04319: CNA27853.1: ALT_INIT.
 DR EMBL: D26562: CNA20773.1: -
 DR EMBL: AE000124: AAC3364.1: -
 DR EMBL: X05810: CNA3958.1: -
 DR EMBL: X05810: CNA3958.1: -
 DR PIR: S07318: S07318.
 DR PIR: S45222: S45222.
 DR EcoGene: EG10303: flnub.
 DR InterPro: IPR000522: FecCD.
 DR Pfam: PF01032: FecCD_family: 2.
 DR ProDom: PD001597: FecCD: 2.
 DR KEGG: C01100: C01100: TranspOcr; Transmembrane; Inner membrane;
 KM Complete proteome: 25
 FT TRANSMEM 5 83 POTENTIAL.
 FT TRANSMEM 93 113 POTENTIAL.
 FT TRANSMEM 118 138 POTENTIAL.
 FT TRANSMEM 147 167 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 240 260 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 FT TRANSMEM 303 323 POTENTIAL.
 FT TRANSMEM 348 368 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 424 444 POTENTIAL.
 FT TRANSMEM 447 467 POTENTIAL.
 FT TRANSMEM 489 509 POTENTIAL.
 FT TRANSMEM 528 548 POTENTIAL.
 FT TRANSMEM 567 587 POTENTIAL.
 FT TRANSMEM 607 627 POTENTIAL.
 FT TRANSMEM 635 655 POTENTIAL.
 FT TRANSMEM 11 11 L -> V (IN REF. 5).
 FT CONFLICT 311 311 MISSING (IN REF. 1 AND 2).
 FT CONFLICT 314 314 MISSING (IN REF. 1 AND 2).
 FT SEQUENCE 660 AA: 70422 MW: 42791578.218096 CIRC64;
 SO SEQUENCE

Query Match 46.08: Score 40: DB 1: Length 660:
 Best Local Similarity 58.38: Pred. No. 1.5e+02:
 Matches 7: Conservative 3: Mismatches 2: Indels 0: Gaps 0:
 Oy 5 VYVLLLVYNGV 16
 Db 127 VVGLVGVAVANG 138

RESULT 42
 1281_BUCAL_BUCAL STANDARD: PRT: 307 AA.
 AC P57358.
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEICAL PROTEIN B0281.
 RN B0281.
 CC Bacterial symbiote (subsp. Acrythosiphon pisum) (Acrythosiphon pisum
 OS Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 NC Symbol: B0281.
 NC NCBL:Taxid:118099.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN: TOKYO 1998.
 KM MEDLINE:20445173: PubMed:10993077:
 RX "Buchnera sp. strain Tokyo 1998, isolated from the endosymbiotic bacterium of aphids
 RL Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE ENAM TRANSPORTER FAMILY. STRONG. TO
 CC S. TYPHIMURUM PAGO.
 CC This SwissProt entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC -----
 DR EMBL: AP001118: BAI12991.1: -
 DR InterPro: IPR001620: DUF6.
 DR Pfam: PF00892: DUF6: 2.
 DR KEGG: C01100: C01100: TranspOcr; Transmembrane; Complete proteome.
 KM Complete proteome: 4
 FT TRANSMEM 4 41 POTENTIAL.
 FT TRANSMEM 41 91 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 146 166 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 267 287 POTENTIAL.
 SO SEQUENCE 307 AA: 35006 MW: 58151AC6F6ACE CIRC64;
 CC

Query Match 45.48: Score 39.5: DB 1: Length 307:
 Best Local Similarity 44.48: Pred. No. 94:
 Matches 8: Conservative 2: Mismatches 3: Indels 5: Gaps 1:
 Oy 3 MLVFL-----LLMGV 15
 Db 4 ILVILFSLVLTWOTT 21

RESULT 43
 1281_BUCAL_BUCAL STANDARD: PRT: 455 AA.
 AC Q92789: Q92789.
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE HEXOSE PHOSPHATE TRANSPORT PROTEIN.
 RN Q92789: Q92789.
 CC Chlamydia pneumoniae (Chlamydia pneumoniae).
 OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 NC NCBL:Taxid:83556.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN: CML023.
 KM MEDLINE:9920606: PubMed:10192388:
 RX "Buchnera sp. strain Tokyo 1998, isolated from the endosymbiotic bacterium of aphids
 RL Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE ENAM TRANSPORTER FAMILY. STRONG. TO
 CC S. TYPHIMURUM PAGO.
 CC This SwissProt entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC or send an email to license@slb.slb.ch).

DR EMBL: X59339; CAA67712.1; -
DR EMBL: X59339; CAA67712.1; -
DR EMBL: Z69123; CAA15515.1; -
DR Subtilist: BG11897; yxjN. -
KW Hypothetical Protein; Transmembrane; Complete proteome.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
SQ SEQUENCE 157 AA; 17382 MW; DDDPFC76A6GCAFP6 CMC64;

Query Match 44.8%; Score 39; DB 1; Length 157;
Best Local Similarity 50.0%; Pred. NO. 65;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Oy 3 MUYVFLILACVT 14
 ::|||||*
Db 30 LITPLELLMGVS 41

Search completed: January 7, 2002, 16:50:36
Job time: 278 sec

DR Prodom: PD002371: Equine_virus_ORF5.1.
 SQ SEQUENCE 255 AA: 28651 MW: 0E8BD432115521F6 CRC64:

Query Match
 Best Local Similarity 61.5%: Pred. No. 29;
 Matches 8: Conservative 3; Mismatches 0: Gaps 0:

OY 1 MSMLVFLLMGV 13
 :|||:|||||
 DB 2 LSNLVFLFLMGV 14

RESULT 12

ID 090KG7 PRELIMINARY: PRT: 255 AA.

AC 090KG7
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DE 090KG7-2001 (TRENBLREL. 17, Last annotation update)
 OS Equine arteritis virus (EAV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage: Nidovirales:
 OC Arteriviridae: Arterivirus.
 OX NCBI_TaxID=11047;
 RX STRAIN=P1;
 RP SEQUENCE FROM N.A.
 RA BALAURUYA U.B.R., HEDGES J.F., NADLER S.A., MCCOLLUM W.H.,
 RA TIMONEY P.J., MACLACHLAN N.J.;
 RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
 RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.";
 RL J. Gen. Virol. 80:1949-1958(1999).
 DR EMBL: AF118772: A054698.1;
 DR InterPro: IPR003241: Arterivirus_ORF5.
 DR EMBL: AF118775: A054716.1;
 DR InterPro: IPR003241: Arterivirus_ORF5.
 DR Pfam: PF00951: Arterivirus_ORF5.1.
 DR Prodom: PD002371: Equine_virus_ORF5.1.
 SQ SEQUENCE 255 AA: 28918 MW: AB92F8FD00D65B CRC64:

Query Match
 Best Local Similarity 54.0%: Score 47; DB 12; Length 255;
 Matches 8: Conservative 3; Mismatches 2; Indels 0: Gaps 0:

OY 1 MSMLVFLLMGV 13
 :|||:|||||
 DB 2 LSNLVFLFLMGV 14

RESULT 13

ID 090KG5 PRELIMINARY: PRT: 255 AA.

AC 090KG5
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE 0905
 OS Equine arteritis virus (EAV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage: Nidovirales:
 OC Arteriviridae: Arterivirus.
 OX NCBI_TaxID=11047;
 RX STRAIN=R2;
 RP SEQUENCE FROM N.A.
 RA BALAURUYA U.B.R., HEDGES J.F., NADLER S.A., MCCOLLUM W.H.,
 RA TIMONEY P.J., MACLACHLAN N.J.;
 RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
 RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.";
 RL J. Gen. Virol. 80:1949-1958(1999).
 DR EMBL: AF118774: A054710.1;
 DR Prodom: PD002371: Equine_virus_ORF5.1.

DR InterPro: IPR001332: Arterivirus_ORF5.
 DR InterPro: IPR003241: Equine_virus_ORF5.
 DR Pfam: PF00951: Arterivirus_ORF5.1.
 DR Prodom: PD002371: Equine_virus_ORF5.1.
 SQ SEQUENCE 255 AA: 28750 MW: D4E8B1CD005C66B CRC64:

Query Match
 Best Local Similarity 54.0%: Score 47; DB 12; Length 255;
 Matches 8: Conservative 3; Mismatches 2; Indels 0: Gaps 0:

OY 1 MSMLVFLLMGV 13
 :|||:|||||
 DB 2 LSNLVFLFLMGV 14

RESULT 14

ID 090KG3 PRELIMINARY: PRT: 255 AA.

AC 090KG3
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DE 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE 0905.
 OS Equine arteritis virus (EAV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage: Nidovirales:
 OC Arteriviridae: Arterivirus.
 OX NCBI_TaxID=11047;
 RX STRAIN=P1;
 RP SEQUENCE FROM N.A.
 RA BALAURUYA U.B.R., HEDGES J.F., NADLER S.A., MCCOLLUM W.H.,
 RA TIMONEY P.J., MACLACHLAN N.J.;
 RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
 RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.";
 RL J. Gen. Virol. 80:1949-1958(1999).
 DR EMBL: AF118775: A054716.1;
 DR InterPro: IPR003241: Arterivirus_ORF5.
 DR InterPro: IPR003241: Arterivirus_ORF5.
 DR Pfam: PF00951: Arterivirus_ORF5.1.
 DR Prodom: PD002371: Equine_virus_ORF5.1.
 SQ SEQUENCE 255 AA: 28767 MW: BB5C0B9525E37AB CRC64:

Query Match
 Best Local Similarity 54.0%: Score 47; DB 12; Length 255;
 Matches 8: Conservative 3; Mismatches 2; Indels 0: Gaps 0:

OY 1 MSMLVFLLMGV 13
 :|||:|||||
 DB 2 LSNLVFLFLMGV 14

RESULT 15

ID 090KG2 PRELIMINARY: PRT: 255 AA.

AC 090KG2
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
 DE 0905
 OS Equine arteritis virus (EAV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage: Nidovirales:
 OC Arteriviridae: Arterivirus.
 OX NCBI_TaxID=11047;
 RX STRAIN=P1;
 RP SEQUENCE FROM N.A.
 RA BALAURUYA U.B.R., HEDGES J.F., NADLER S.A., MCCOLLUM W.H.,
 RA TIMONEY P.J., MACLACHLAN N.J.;
 RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
 RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.";
 RL J. Gen. Virol. 80:1949-1958(1999).
 DR EMBL: AF118774: A054710.1;
 DR Prodom: PD002371: Equine_virus_ORF5.1.

```

RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.*
RT J. Gen. Virol. 80:1949-1958(1999).
DR EMBL: AF118776: AAD54722.1; -.
RA InterPro: IPR001332: Artercl_glycop.
DR InterPro: IPR003241: Equine_virus_ORF5.
RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.*
DR EMBL: AF118776: AAD54722.1; -.
DR EMBL: AF118773: AAD54758.1; -.
DR EMBL: AF118781: AAD54752.1; -.
DR InterPro: IPR001332: Artercl_glycop.
DR InterPro: IPR003241: Equine_virus_ORF5.
DR Pfam: PF00951: Artercl_glycop.1.
DR Prodom: PD002371: Equine_virus_ORF5; 1.
SQ SEQUENCE 255 AA: 28798 MW: 4796CALIETRI16F3 CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Prod. No. 29;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
DB 2 LSMIVLFLMGV 14

RESULT 17
O9PMV4 PRELIMINARY; PRT; 255 AA.
AC O9PMV4: 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
DE ORF5
DR Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CA3, AT, AND A2;
RC EMBL: AF118776: AAD54692.1; -.
RA Balasuriya U.B.R., Hodges J.F., Nadler S.A., McCollum W.H.,
RA Timoney P.J., MacLachlan N.J.;
RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.*"
RL J. Gen. Virol. 80:1949-1958(1999).
DR EMBL: AF118771: AAD54692.1; -.
DR EMBL: AF118770: AAD54690.1; -.
DR EMBL: AF118772: AAD54691.1; -.
DR InterPro: IPR001332: Artercl_glycop.
DR InterPro: IPR003241: Equine_virus_ORF5.
DR Pfam: PF00951: Artercl_glycop.1.
DR Prodom: PD002371: Equine_virus_ORF5; 1.
SQ SEQUENCE 255 AA: 28739 MW: 9AN6A5270A1E1BCA CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Prod. No. 29;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
DB 2 LSMIVLFLMGV 14

```

```

RC STRAIN=RO, R1, AND BT-PAGE;
RC MEDLINE:99394665: PubMed-10466790;
RA Balasuriya U.B.R., Hodges J.F., Nadler S.A., McCollum W.H.,
RA Timoney P.J., MacLachlan N.J.;
RT "Genetic Stability of Equine Arteritis Virus during Horizontal and
RT Vertical Transmission in an Outbreak of Equine Viral Arteritis.*"
DR EMBL: AF118776: AAD54758.1; -.
DR EMBL: AF118773: AAD54758.1; -.
DR EMBL: AF118781: AAD54752.1; -.
DR InterPro: IPR001332: Artercl_glycop.
DR InterPro: IPR003241: Equine_virus_ORF5.
DR Pfam: PF00951: Artercl_glycop.1.
DR Prodom: PD002371: Equine_virus_ORF5; 1.
SQ SEQUENCE 255 AA: 28725 MW: DB72F0173BAAC CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Prod. No. 29;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
DB 2 LSMIVLFLMGV 14

RESULT 18
Q66483 Q66483 PRELIMINARY; PRT; 255 AA.
AC Q66483: 09YUN3.
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLREL. 17, Last annotation update)
DE EAV CAN6 ORF5 (FIELD STRAIN CAN6 ORF5 5, 6 AND 7 GENES, COMPLETE
DE CDS)
DR Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EAV CAN6 (FIELD ISOLATE);
RC EMBL: AF118776: AAD54692.1; -.
RA Balasuriya U.B.R., Hodges J.F., Patton J.F., Timoney P.J.;
RT "Sequence Analysis and Phylogenetic Characterization of Two Strains of
RT Equine Arteritis Virus Isolated from Separately Infected Stallions
RT on the Same Ranch in California.*"
RL J. Virol. 69:0-0(1996).
DR EMBL: AF118776: AAD54692.1; -.
DR EMBL: AF118773: AAD54758.1; -.
DR EMBL: AF118781: AAD54752.1; -.
DR InterPro: IPR001332: Artercl_glycop.
DR InterPro: IPR003241: Equine_virus_ORF5.
DR Pfam: PF00951: Artercl_glycop.1.
DR Prodom: PD002371: Equine_virus_ORF5; 1.
SQ SEQUENCE 255 AA: 8D0685A012198E02 CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Prod. No. 29;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
DB 2 LSMIVLFLMGV 14

```

```

RESULT 19
OQ995W0: PRELIMINARY: PRT: 255 AA.
AC 0995W0:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GL PROTEIN.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC NCBI_TaxID:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9901008;
RA Larsen L.E., Storgaard T., Holm E.;
RT Phylogenetic characterization of the GL protein of equine arteritis
RT virus isolated from semen of asymptotic stallions and fatal cases of
RT equine viral arteritis.
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF247547; AAK37630.1; -
SQ SEQUENCE 255 AA: 28760 MW: 550E905C4960DE4 CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 29;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
Db 2 LSMIVLFLMGV 14

RESULT 20
OQ995W0: PRELIMINARY: PRT: 255 AA.
AC 0995W0:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GL PROTEIN.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OC NCBI_TaxID:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9910583;
RA Larsen L.E., Storgaard T., Holm E.;
RT Phylogenetic characterization of the GL protein of equine arteritis
RT virus isolated from semen of asymptotic stallions and fatal cases of
RT equine viral arteritis.
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF247548; AAK37631.1; -
SQ SEQUENCE 255 AA: 28810 MW: 10089E2056827CC2 CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 29;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
Db 2 LSMIVLFLMGV 14

RESULT 21
OQ995V9: PRELIMINARY: PRT: 255 AA.
AC 0995V9:
DT 01-JUN-2001 (TEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GL PROTEIN.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OC NCBI_TaxID:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9910583;
RA Larsen L.E., Storgaard T., Holm E.;
RT Phylogenetic characterization of the GL protein of equine arteritis
RT virus isolated from semen of asymptotic stallions and fatal cases of
RT equine viral arteritis.
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF247549; AAK37632.1; -
SQ SEQUENCE 255 AA: 28825 MW: 51DC0C0625A72BE4 CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 29;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
Db 2 LSMIVLFLMGV 14

RESULT 22
OQ99Y16: PRELIMINARY: PRT: 255 AA.
AC 099Y16:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GL PROTEIN.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OC NCBI_TaxID:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2344-9; AND 2344-8;
RA Larsen L.E., Storgaard T., Holm E.;
RT Phylogenetic characterization of the GL protein of equine arteritis
RT virus isolated from semen of asymptotic stallions and fatal cases of
RT equine viral arteritis.
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF247541; AAK37624.1; -
SQ SEQUENCE 255 AA: 28783 MW: 6B5DC0CB79512EFF CRC64;

Query Match 54.0%: Score 47; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 29;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 MSMLVFFLLMGV 13
Db 2 LSMIVLFLMGV 14

RESULT 23
OQ9HXB6: PRELIMINARY: PRT: 300 AA.
ID 09HXB6:
AC 09HXB6:
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PA3897.
CN PA3897.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

```

```

OC Pseudomonas.
NCBI_TaxID:287;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE:20437337; PubMed:10984043;
RA Slocver C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hittinge W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Brinkman F.S.L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Spiller S.N., Folger K.R., Kras A., Lashby K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA Kaper E.B., et al. 1998. The genome of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen. Nature 391:392-398.
RL Nature 406:959-964(2000).
DR EMBL: AE004807; AAC07284.1; -.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 300 AA; 32291 MW; 390033DEC3159a1c CRC64;

Query Match 54.0%; Score 47; DB 2; Length 300;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Oy 1 MSLLVFLLMGVTG 15
DB 3 LSLVLTWLLMGVTG 17

RESULT 24
Oy09C17 PRELIMINARY; PRT; 389 AA.
ID 09C17
RA 09C17
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE TRANSPORT PROTEIN.
OS XPD281.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC NCBI_TaxID:2371;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE:9A5C;
RA Slocver C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hittinge W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Brinkman F.S.L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Spiller S.N., Folger K.R., Kras A., Lashby K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA Kaper E.B., et al. 1998. The genome of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen. Nature 391:392-398.
RL Nature 406:959-964(2000).
DR EMBL: AE004807; AAC07284.1; -.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 300 AA; 32291 MW; 390033DEC3159a1c CRC64;

```

```

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RA "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
RA EMBL: AE001881; AAP81094.1; -.
RA InterPro: IPR002589; IPR0110.
RA NCBI_TaxID:287;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE:20437337; PubMed:10984043;
RA Slocver C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hittinge W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Brinkman F.S.L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Spiller S.N., Folger K.R., Kras A., Lashby K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA Kaper E.B., et al. 1998. The genome of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen. Nature 391:392-398.
RL Nature 406:959-964(2000).
DR EMBL: AE004807; AAC07284.1; -.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 300 AA; 32291 MW; 390033DEC3159a1c CRC64;

Query Match 52.9%; Score 46; DB 6; Length 529;
Best Local Similarity 56.2%; Pred. No. 73;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 MSLLVFLLMGVTG 16
DB 8 VSAVAFLLMGVTG 23

RESULT 26
Oy14F8 PRELIMINARY; PRT; 448 AA.
ID 0914F8
RA 0914F8
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-JUN-2001 (TEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE TWO-COMPONENT SENSOR PHOQ.
OS PHOQ OR PAI180.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
NCBI_TaxID:287;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE:20437337; PubMed:10984043;
RA Slocver C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hittinge W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Brinkman F.S.L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Spiller S.N., Folger K.R., Kras A., Lashby K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA Kaper E.B., et al. 1998. The genome of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen. Nature 391:392-398.
RL Nature 406:959-964(2000).
DR EMBL: AE004807; AAC07284.1; -.
DR InterPro: IPR000620; DUF6.
DR Pfam: PF00892; DUF6; 2.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 529 AA; 60678 MW; 10EBCD920A415F76 CRC64;

```

```

RA Brody L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizner J., Saiter M.H., Hancock R.E.W., Loty S., Olson M.V.,
RA Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RA Escherichia coli O157:H7 strain (GenBank accession: F003962200).
CC 10-11-2000.
CC 11-11-2000.
CC 12-11-2000.
CC 13-11-2000.
CC 14-11-2000.
CC 15-11-2000.
CC 16-11-2000.
CC 17-11-2000.
CC 18-11-2000.
CC 19-11-2000.
CC 20-11-2000.
CC 21-11-2000.
CC 22-11-2000.
CC 23-11-2000.
CC 24-11-2000.
CC 25-11-2000.
CC 26-11-2000.
CC 27-11-2000.
CC 28-11-2000.
CC 29-11-2000.
CC 30-11-2000.
CC 31-11-2000.
CC 32-11-2000.
CC 33-11-2000.
CC 34-11-2000.
CC 35-11-2000.
CC 36-11-2000.
CC 37-11-2000.
CC 38-11-2000.
CC 39-11-2000.
CC 40-11-2000.
CC 41-11-2000.
CC 42-11-2000.
CC 43-11-2000.
CC 44-11-2000.
CC 45-11-2000.
CC 46-11-2000.
CC 47-11-2000.
CC 48-11-2000.
CC 49-11-2000.
CC 50-11-2000.
CC 51-11-2000.
CC 52-11-2000.
CC 53-11-2000.
CC 54-11-2000.
CC 55-11-2000.
CC 56-11-2000.
CC 57-11-2000.
CC 58-11-2000.
CC 59-11-2000.
CC 60-11-2000.
CC 61-11-2000.
CC 62-11-2000.
CC 63-11-2000.
CC 64-11-2000.
CC 65-11-2000.
CC 66-11-2000.
CC 67-11-2000.
CC 68-11-2000.
CC 69-11-2000.
CC 70-11-2000.
CC 71-11-2000.
CC 72-11-2000.
CC 73-11-2000.
CC 74-11-2000.
CC 75-11-2000.
CC 76-11-2000.
CC 77-11-2000.
CC 78-11-2000.
CC 79-11-2000.
CC 80-11-2000.
CC 81-11-2000.
CC 82-11-2000.
CC 83-11-2000.
CC 84-11-2000.
CC 85-11-2000.
CC 86-11-2000.
CC 87-11-2000.
CC 88-11-2000.
CC 89-11-2000.
CC 90-11-2000.
CC 91-11-2000.
CC 92-11-2000.
CC 93-11-2000.
CC 94-11-2000.
CC 95-11-2000.
CC 96-11-2000.
CC 97-11-2000.
CC 98-11-2000.
CC 99-11-2000.
CC 100-11-2000.
CC 101-11-2000.
CC 102-11-2000.
CC 103-11-2000.
CC 104-11-2000.
CC 105-11-2000.
CC 106-11-2000.
CC 107-11-2000.
CC 108-11-2000.
CC 109-11-2000.
CC 110-11-2000.
CC 111-11-2000.
CC 112-11-2000.
CC 113-11-2000.
CC 114-11-2000.
CC 115-11-2000.
CC 116-11-2000.
CC 117-11-2000.
CC 118-11-2000.
CC 119-11-2000.
CC 120-11-2000.
CC 121-11-2000.
CC 122-11-2000.
CC 123-11-2000.
CC 124-11-2000.
CC 125-11-2000.
CC 126-11-2000.
CC 127-11-2000.
CC 128-11-2000.
CC 129-11-2000.
CC 130-11-2000.
CC 131-11-2000.
CC 132-11-2000.
CC 133-11-2000.
CC 134-11-2000.
CC 135-11-2000.
CC 136-11-2000.
CC 137-11-2000.
CC 138-11-2000.
CC 139-11-2000.
CC 140-11-2000.
CC 141-11-2000.
CC 142-11-2000.
CC 143-11-2000.
CC 144-11-2000.
CC 145-11-2000.
CC 146-11-2000.
CC 147-11-2000.
CC 148-11-2000.
CC 149-11-2000.
CC 150-11-2000.
CC 151-11-2000.
CC 152-11-2000.
CC 153-11-2000.
CC 154-11-2000.
CC 155-11-2000.
CC 156-11-2000.
CC 157-11-2000.
CC 158-11-2000.
CC 159-11-2000.
CC 160-11-2000.
CC 161-11-2000.
CC 162-11-2000.
CC 163-11-2000.
CC 164-11-2000.
CC 165-11-2000.
CC 166-11-2000.
CC 167-11-2000.
CC 168-11-2000.
CC 169-11-2000.
CC 170-11-2000.
CC 171-11-2000.
CC 172-11-2000.
CC 173-11-2000.
CC 174-11-2000.
CC 175-11-2000.
CC 176-11-2000.
CC 177-11-2000.
CC 178-11-2000.
CC 179-11-2000.
CC 180-11-2000.
CC 181-11-2000.
CC 182-11-2000.
CC 183-11-2000.
CC 184-11-2000.
CC 185-11-2000.
CC 186-11-2000.
CC 187-11-2000.
CC 188-11-2000.
CC 189-11-2000.
CC 190-11-2000.
CC 191-11-2000.
CC 192-11-2000.
CC 193-11-2000.
CC 194-11-2000.
CC 195-11-2000.
CC 196-11-2000.
CC 197-11-2000.
CC 198-11-2000.
CC 199-11-2000.
CC 200-11-2000.
CC 201-11-2000.
CC 202-11-2000.
CC 203-11-2000.
CC 204-11-2000.
CC 205-11-2000.
CC 206-11-2000.
CC 207-11-2000.
CC 208-11-2000.
CC 209-11-2000.
CC 210-11-2000.
CC 211-11-2000.
CC 212-11-2000.
CC 213-11-2000.
CC 214-11-2000.
CC 215-11-2000.
CC 216-11-2000.
CC 217-11-2000.
CC 218-11-2000.
CC 219-11-2000.
CC 220-11-2000.
CC 221-11-2000.
CC 222-11-2000.
CC 223-11-2000.
CC 224-11-2000.
CC 225-11-2000.
CC 226-11-2000.
CC 227-11-2000.
CC 228-11-2000.
CC 229-11-2000.
CC 230-11-2000.
CC 231-11-2000.
CC 232-11-2000.
CC 233-11-2000.
CC 234-11-2000.
CC 235-11-2000.
CC 236-11-2000.
CC 237-11-2000.
CC 238-11-2000.
CC 239-11-2000.
CC 240-11-2000.
CC 241-11-2000.
CC 242-11-2000.
CC 243-11-2000.
CC 244-11-2000.
CC 245-11-2000.
CC 246-11-2000.
CC 247-11-2000.
CC 248-11-2000.
CC 249-11-2000.
CC 250-11-2000.
CC 251-11-2000.
CC 252-11-2000.
CC 253-11-2000.
CC 254-11-2000.
CC 255-11-2000.
CC 256-11-2000.
CC 257-11-2000.
CC 258-11-2000.
CC 259-11-2000.
CC 260-11-2000.
CC 261-11-2000.
CC 262-11-2000.
CC 263-11-2000.
CC 264-11-2000.
CC 265-11-2000.
CC 266-11-2000.
CC 267-11-2000.
CC 268-11-2000.
CC 269-11-2000.
CC 270-11-2000.
CC 271-11-2000.
CC 272-11-2000.
CC 273-11-2000.
CC 274-11-2000.
CC 275-11-2000.
CC 276-11-2000.
CC 277-11-2000.
CC 278-11-2000.
CC 279-11-2000
```

DT	01-MAY-1997	(TEMBLrel_03, Last sequence update)
DT	01-JUN-2001	(TEMBLrel_17, Last annotation update)
DE	CYVI (FRAGMENT).	
DN	CYVI.	
OS	Pseudomonas fluorescens.	
OC	Pseudomonas fluorescencia; gamma subdivision; Pseudomonadaceae.	
NCBI_TaxID=294;		
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN-GH-1;	
RA	Vitold J.P.M., Brouwers G.J., Corsjens P.L.A.M., Dulk J.,	
RL	Virind de Jong E.N.; to the EMBL/GenBank/DDJB databases.	
DR	EMBL:U85716; AF065583	
DI	Interpro: IPR002541; CYC.asm.	
DR	Interpro: IPR003557; CYC-biog_CmcC.	
PLam:	PF01578; CYC.asm. 1.	
DR	PRINTS: PR01365; COMBIGNSTS.	
FT	NON_TER	177
SQ	SEQUENCE	177 AA: TCDP9BF0B7E58Z44 CRC64:
Query Match		
Best Local Similarity	51.7%:	Score 45: DB 2: Length 177:
Matches	8: Conservative	2: Mismatches
Oy	5 VPELLMGVYC 16	
Dy	8 LVLLNGVING 19	
RESULT 29		
ID	P87643	PRELIMINARY: PRT: 255 AA.
AC	P87643	
DT	01-MAY-1997	(TEMBLrel_03, Created)
DT	01-JUN-2001	(TEMBLrel_17, Last sequence update)
DT	01-JUN-2001	(TEMBLrel_17, Last annotation update)
DE	STRAIN VIENNA ORFS.	
OS	Equine arteritis virus (EAV).	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Nitovirales;	
OC	Arteriviridae; Arterivirus.	
DX	NCBI_TaxID=11047;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN-VIENNA:	
RA	Archambault D., St-Laurent G., Lepage N.;	
RL	Submitted (JAN-1996) to the EMBL/GenBank/DDJB databases.	
DR	EMBL: U46952; AAB50283.1;	
DI	Interpro: IPR001332; Arteriv_glycop.	
DR	Interpro: IPR003241; Equine_virus_ORFS.	
PLam:	PD002371; Arteriv_glycop. 1 ORFS: 1	
DR	PRINTS: PD002371;	
SQ	SEQUENCE	255 AA: 2R635 MW: 563DADB17035D5 CRC64:
Query Match		
Best Local Similarity	51.7%:	Score 45: DB 12: Length 255:
Matches	8: Conservative	3: Mismatches
Oy	1 MSNVVLFLVCV 13	
Dy	2 LSMVLSLMCV 14	
RESULT 30		
ID	P066481	PRELIMINARY: PRT: 255 AA.
AC	P066481	
DT	01-NOV-1996	(TEMBLrel_01, Created)
DT	01-NOV-1996	(TEMBLrel_01, Last sequence update)
DT	01-JUN-2001	(TEMBLrel_17, Last annotation update)

```

DE EAV_AUT68. ORF5.
OC Equine arteritis virus (EAV).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Nidovirales:
OC Arteriviridae: Arterivirus.
RN Medl:taxid:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EAV AUT68 (FIELD ISOLATE);
RC MEDLINE:96130216; PubMed:8553578;
RA Balasuriya U.B., Timoney P.J., McCollum W.H., MacLachlan N.J.;
RA "Phylogenetic analysis of open reading frame 5 of field isolates of
RA equine arteritis virus (EAV) and identification of conserved and
RA nonconserved regions in the GL envelope glycoprotein.";
RL Virology 214:690-697(1995).
DE EMBL: U38592; AA92897.1; -.
DR InterPro: IPR001332; ArterL_glycop.
DR InterPro: IPR003241; Equine_virus_ORF5.
DR Pfam: PF00951; ArterL_glycop.1.
DR ProDom: PD002371; Equine_virus_ORF5.1.
SO SEQUENCE 255 AA; 28550 MW; 505D250187703D5 CRC64;

Query Match 51.7%: Score 45; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 54;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NSMLVFWLLMAY 13
DB 2 LSNVLTSLMAY 14

RESULT 31
O66482 PRELIMINARY: PRT; 255 AA.
ID O66482.
DT 01-NOV-1996 (TRMBLrel. 01; Created)
DT 01-NOV-1996 (TRMBLrel. 01; Last sequence update)
DT 01-JUN-2001 (TRMBLrel. 17; Last annotation update)
DE EAV SW64. ORF5.
DE Equine arteritis virus (EAV).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Nidovirales:
OC Arteriviridae: Arterivirus.
RN Medl:taxid:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EAV SW64 (FIELD ISOLATES);
RC MEDLINE:96130216; PubMed:8553578;
RA Balasuriya U.B., Timoney P.J., McCollum W.H., MacLachlan N.J.;
RA "Phylogenetic analysis of open reading frame 5 of field isolates of
RA equine arteritis virus (EAV) and identification of conserved and
RA nonconserved regions in the GL envelope glycoprotein.";
RL Virology 214:690-697(1995).
DE EMBL: U38609; AA92914.1; -.
DR InterPro: IPR001332; ArterL_glycop.
DR InterPro: IPR003241; Equine_virus_ORF5.
DR Pfam: PF00951; ArterL_glycop.1.
DR ProDom: PD002371; Equine_virus_ORF5.1.
SO SEQUENCE 255 AA; 28649 MW; DEC7A9483808C10 CRC64;

Query Match 51.7%: Score 45; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 54;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NSMLVFWLLMAY 13
DB 2 LSNVLTSLMAY 14

RESULT 32
O66482 PRELIMINARY: PRT; 255 AA.
ID O66482.
DT 01-NOV-1996 (TRMBLrel. 01; Created)
DT 01-NOV-1996 (TRMBLrel. 01; Last sequence update)
DT 01-JUN-2001 (TRMBLrel. 17; Last annotation update)
DE EAV SW64. ORF5.
DE Equine arteritis virus (EAV).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Nidovirales:
OC Arteriviridae: Arterivirus.
RN Medl:taxid:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EAV SW64 (FIELD ISOLATES);
RC MEDLINE:96130216; PubMed:8553578;
RA Balasuriya U.B., Timoney P.J., McCollum W.H., MacLachlan N.J.;
RA "Phylogenetic analysis of open reading frame 5 of field isolates of
RA equine arteritis virus (EAV) and identification of conserved and
RA nonconserved regions in the GL envelope glycoprotein.";
RL Virology 214:690-697(1995).
DE EMBL: U38609; AA92914.1; -.
DR InterPro: IPR001332; ArterL_glycop.
DR InterPro: IPR003241; Equine_virus_ORF5.
DR Pfam: PF00951; ArterL_glycop.1.
DR ProDom: PD002371; Equine_virus_ORF5.1.
SO SEQUENCE 255 AA; 28649 MW; DEC7A9483808C10 CRC64;

```

```

DT 01-NOV-1996 (TRMBLrel. 01; Created)
DT 01-NOV-1996 (TRMBLrel. 01; Last sequence update)
DT 01-JUN-2001 (TRMBLrel. 17; Last annotation update)
DE EAV A267. ORF5.
DE Equine arteritis virus (EAV).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Nidovirales:
OC Arteriviridae: Arterivirus.
RN Medl:taxid:11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EAV A267 (FIELD ISOLATE);
RC MEDLINE:96086337; PubMed:9838958;
RA Balasuriya U.B., Evermann J.F., Hedges J.F., McKelrnan A.J.,
RA Mitten J.O., Beyer J.C., McCollum W.H., Timoney P.J., MacLachlan N.J.;
RA "Serologic and molecular characterization of an abortifacient strain of
RA equine arteritis virus isolated from infective frozen semen and an
RA aborted equine fetus.";
RL J. Am. Vet. Med. Assoc. 213:1586-1589(1998).
DE EMBL: AF065824; AA013738.1; -.
DR InterPro: IPR001332; ArterL_glycop.
DR InterPro: IPR003241; Equine_virus_ORF5.
DR Pfam: PF00951; ArterL_glycop.1.
DR ProDom: PD002371; Equine_virus_ORF5.1.
DR Hypothetical Protein.
SO SEQUENCE 255 AA; 28697 MW; 290CFE42341FFD67 CRC64;

Query Match 51.7%: Score 45; DB 12; Length 255;
Best Local Similarity 61.5%: Pred. No. 54;
Matches 8: Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NSMLVFWLLMAY 13
DB 2 LSNVLTSLMAY 14

RESULT 33
O9RXV3 PRELIMINARY: PRT; 275 AA.
ID O9RXV3.
DT 01-MAY-2000 (TRMBLrel. 13; Created)
DT 01-MAY-2000 (TRMBLrel. 13; Last sequence update)
DT 01-MAR-2001 (TRMBLrel. 16; Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
DE DR0203.
OC Deinococcus radiodurans.
OC Bacteria: Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN Medl:taxid:1295;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RI;
RC MEDLINE:20036896; PubMed:10567266;
RA White O., Eissen J.A., Heideberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hart D.H., Gaim M.L., Nelson W.C., Richardson D.L.,
RA Motil K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Manoharven S.J., Lam J., McInerney J., Zaleska C.,
RA McKelvey J., Heideberg J., Mitterback T., Zaleska C.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
RA radiodurans RI.";
RL Science 286:1571-1577(1999).
DE EMBL: AE001882; AA09708.1; -.

```


RESULT 44

OSKFKT 44 PRELIMINARY: PRT: 417 AA.

AC 09EPY7.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE BH0472 PROTEIN.

GN BH0472.

OS Bacillus halodurans.

OC Bacillus/Enterobacteriaceae; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_Taxid:86665;

RN [1]_Taxid:86665;

RP SEQUENCE FROM N.A.

RC STRAIN:G-125 / JCM 9153;

RE MEDLINE:20512582; PubMed:1058132;

RA TAKAMI H., NAKASONE K., TAKAKI Y., MENO G., SASAKI R., MASUI N.,

RI HORIKAWA K., NAKAMURA Y., OGAWARA N., KUNIDA S.,

RT "Complete genome sequence of the alkaliophilic bacterium *Bacillus*

halodurans and genomic sequence comparison with *Bacillus subtilis*."

RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL: AF001508; BAB04391.1; "

DR InterPro: IPR001958; ICR-TECA.

DR SwissProt: P60033; CHTPRA.

KM Complete protein.

SO SEQUENCE 417 AA; 45439 MW; 17C32402CB1E4M9 CRC64;

Query Match

Best Local Similarity 50.6%; Score 44; DB 2; Length 417;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 VVFLLMGVTV 15

DB 104 VITRLKMGIM 114

RESULT 45

OSKFKT 45 PRELIMINARY: PRT: 471 AA.

AC 09EPY6.

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE PUTATIVE UNDECAPRENTL-PHOSPHATE GALACTOSEPHOSPHOTRANSFERASE.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Actinobacterium

OX NCBI_Taxid:715;

RN [1]_Taxid:715;

RP SEQUENCE FROM N.A.

RC STRAIN:4074;

RE LARUE J., RLOUX S., MADE M.M., CHAMPLIN F.R., HOLMAN S.C.,

RT "Identification of genes involved in the pathogenicity of *Bacillus*

pleuropneumoniae serotype 1 O-antigen and biological properties of

rough mutants."

RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF329452; AAG45943.1; "

DR InterPro: IPR003362; BacL_transf.

DR SwissProt: P60033; CHTPRA.

DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.

SO SEQUENCE 471 AA; 55192 MW; C5B445ADB7B3513 CRC64;

OY 10 LNCVTWC 16
DB 110 LNWLTWG 116

Search completed: January 7, 2002, 16:51:30
Job time: 277 sec

Query Match 50.6%; Score 44; DB 2; Length 471;
Best Local Similarity 85.7%; Pred. No. 1,2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0;

REFERENCE 1 (sites)
 AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigaeta,K., Senba,T.,
 Metsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
 Fujimura,T., Ono,T., Yamada,K., Fujii,T., Ozaki,K., Hiro,M.,
 Tanaka,T., Nakamura,Y., Isogai,T., Shishihara,T.,
 NEDO human cDNA sequencing project
 TITLE 2 (bases 1 to 2175)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2175)
 AUTHORS Shigaeta,S., Suzuki,Y., Ota,T., Odayashii,M., Nishi,T., Isogai,T.,
 Shishihara,T., Tanaka,T., and Nakamura,Y.
 JOURNAL Submitted (28-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sumio
 Sugano, Institute of Medical Science, University of Tokyo,
 Laboratory of Genome Structure Analysis, Human Genome Center;
 Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8633, Japan
 (E-mail:odaia@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Health, Labour and Welfare, Japan. This is a full-length cDNA library
 sequencing. Research Association for Biotechnology, DNA
 Technology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source 1. 2175
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1ld="1NC1"
 /clone_1ld="1NC1"
 /tissue_type="human lung"
 /note="Cloning vector: pMCS8PL3"
 BASE COUNT 520 a 598 c 626 g 431 t
 ORIGIN
 Alignment_scores:
 Quality: 68.50 Length: 17
 Ratio: 4.281 Gaps: 1
 Percent Similarity: 94.118 Percent Identity: 82.353
 alignment_block:
 US-09-471-276-831_COPY_1_16 x AK021222/rev ..
 Align seq 1/1 to reverse of: AK021222 from: 1 to: 2175
 1 MetSerMetLeuValValPheLeuLeuLeuTTrGly...ValThrTrpCl 16
 |
 1134 ATGTCATGCTGCTGCTGCTTCTTCTGCTGCTGCTGAGGCTCTGCGG 1085
 |
 16 y 16
 1084 T 1084
 seq_name: gb_hlg:AC012313
 seq_documentation_block:
 LOCUS AC012313.1186451 bp DNA HTG 05-AUG-2001
 DEFINITION Homo sapiens chromosome 19 clone CTD-2619J3, *** SEQUENCING IN
 PROGRESS ***, 2 ordered pieces.
 ACCESSION AC012313
 VERSION AC012313.6 GI:15145671
 KEYWORDS HTG: HTGS_PHA52; HTGS_ACTIVEIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 186451)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE DOE Joint Submissions
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 186451)

AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (22-OCT-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT
 NOTE: This is a working draft.
 * consists of 2 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as it is available and
 * the accession number will be preserved
 * 1 181942: contig of 181942 bp in length
 * 181943 182042: gap of unknown length
 * 182043 186451: contig of 4409 bp in length.
 *****Genome Center
 Center: Joint Genome Institute
 Web site: http://www.jgi.doe.gov
 Project Information
 Center Project Name: 835906
 Center clone name: CTD-2619J13

 Summary Statistics
 Consensus quality: 184784 bases at least Q40
 Consensus quality: 183394 bases at least Q30
 Consensus quality: 185487 bases at least Q20
 Estimated insert size: 184270; agarose-fp estimation
 Estimated insert size: 185541; sum-of-contigs
 estimation
 Quality coverage: 0.95 in Q20 bases; agarose-fp
 estimation
 Quality coverage: 1 in Q20 bases; sum-of-contigs
 estimation
 Location/Qualifiers
 source 1. 186451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19p13.3"
 /clone_CTD="2619J3"
 BASE COUNT 44285 a 49150 c 49575 g 43301 t 100 others
 ORIGIN
 Alignment_scores:
 Quality: 68.50 Length: 17
 Ratio: 4.281 Gaps: 1
 Percent Similarity: 94.118 Percent Identity: 82.353
 alignment_block:
 US-09-471-276-831_COPY_1_16 x AC012313/rev ..
 Align seq 1/1 to reverse of: AC012313 from: 1 to: 186451
 1 MetSerMetLeuValValPheLeuLeuLeuTTrGly...ValThrTrpCl 16
 |
 11643 ATGTCATGCTGCTGCTGCTTCTTCTGCTGCTGCTGAGGCTCTGCGG 11594
 |
 16 y 16
 11593 T 11593
 seq_name: gb_hlg:AC022955
 seq_documentation_block:
 LOCUS AC022955.159560 bp DNA HTG 16-MAR-2000
 DEFINITION Homo sapiens clone Rpi1-15310, WORKING DRAFT SEQUENCE, 44
 unordered pieces.
 ACCESSION AC022955
 VERSION AC022955.3 GI:7249299

[illegible]

9730	9829:	gap of	100 bp	In Length
9830	10597:	contig of 1166 bp	In Length	
10998	11097:	gap of	100 bp	In Length
11098	12160:	contig of 1504 bp	In Length	
12602	12701:	gap of	100 bp	In Length
14322	144131:	contig of 1630 bp	In Length	
14322	15960:	contig of 1529 bp	In Length	
15961	16060:	gap of	100 bp	In Length
16061	17698:	contig of 1638 bp	In Length	
17699	17798:	gap of	100 bp	In Length
20096	20195:	contig of 2297 bp	In Length	
20096	20195:	gap of	100 bp	In Length
22181	22280:	gap of	100 bp	In Length
22281	24180:	contig of 1840 bp	In Length	
24121	24220:	gap of	100 bp	In Length
24221	26233:	contig of 2013 bp	In Length	
26234	26333:	gap of	100 bp	In Length
26334	28235:	contig of 1902 bp	In Length	
28235	28333:	gap of	100 bp	In Length
30683	30965:	gap of	100 bp	In Length
30963	34201:	contig of 3239 bp	In Length	
34202	34301:	gap of	100 bp	In Length
34302	37474:	contig of 3173 bp	In Length	
37475	37574:	gap of	100 bp	In Length
37575	40724:	contig of 3150 bp	In Length	
40725	40823:	gap of	100 bp	In Length
43316	43415:	gap of	100 bp	In Length
43416	46376:	contig of 3331 bp	In Length	
46377	46866:	gap of	100 bp	In Length
46867	46873:	contig of 3027 bp	In Length	
46874	49973:	gap of	100 bp	In Length
49974	53197:	contig of 3324 bp	In Length	
53198	55630:	gap of	100 bp	In Length
55631	56460:	gap of	100 bp	In Length
56461	60731:	contig of 4291 bp	In Length	
60732	60851:	gap of	100 bp	In Length
60852	65004:	contig of 4153 bp	In Length	
65005	65104:	gap of	100 bp	In Length
65105	66532:	contig of 4418 bp	In Length	
66533	72326:	contig of 3847 bp	In Length	
72327	73389:	gap of	100 bp	In Length
73390	76613:	contig of 3324 bp	In Length	
76614	76713:	gap of	100 bp	In Length
76714	81133:	contig of 4440 bp	In Length	
81134	81535:	gap of	100 bp	In Length
81536	86590:	contig of 1667 bp	In Length	
86591	91545:	contig of 5525 bp	In Length	
91546	91645:	gap of	100 bp	In Length
91646	96877:	contig of 5232 bp	In Length	
96878	96977:	gap of	100 bp	In Length
102219	102293:	contig of 5155 bp	In Length	
102294	102299:	gap of	100 bp	In Length
102552	107650:	contig of 5032 bp	In Length	
107652	112462:	contig of 4791 bp	In Length	
112463	112552:	gap of	100 bp	In Length
112553	117674:	contig of 5122 bp	In Length	
117675	117774:	gap of	100 bp	In Length
117775	125917:	contig of 8143 bp	In Length	
125918	126017:	gap of	100 bp	In Length
126019	132844:	contig of 100 bp	In Length	
132850	137619:	contig of 5235 bp	In Length	
137620	137719:	gap of	100 bp	In Length
137720	144448:	contig of 6779 bp	In Length	
144449	144599:	gap of	100 bp	In Length


```

misc_feature      84480..38175
/note="assembly_fragment"
misc_feature      98276..112174
/note="assembly_fragment"
misc_feature      112225..126338
/note="assembly_fragment"
misc_feature      126339..143537
vector_side:right*
misc_feature      143538..159722
/note="assembly_fragment"
misc_feature      159723..163371
/note="assembly_fragment"
misc_feature      163372..169679
/note="assembly_fragment"
misc_feature      169680..175983
/note="assembly_fragment"
misc_feature      175984..182287
/note="assembly_fragment"
misc_feature      182288..188591
/note="assembly_fragment"
misc_feature      188592..194895
/note="assembly_fragment"
misc_feature      194896..201199
/note="assembly_fragment"
misc_feature      201200..207503
/note="assembly_fragment"
misc_feature      207504..213807
/note="assembly_fragment"
misc_feature      213808..220111
/note="assembly_fragment"
misc_feature      220112..226415
/note="assembly_fragment"
misc_feature      226416..232719
/note="assembly_fragment"
misc_feature      232720..239023
/note="assembly_fragment"
misc_feature      239024..245327
/note="assembly_fragment"
misc_feature      245328..251631
/note="assembly_fragment"
misc_feature      251632..257935
/note="assembly_fragment"
misc_feature      257936..264239
/note="assembly_fragment"
misc_feature      264240..270543
/note="assembly_fragment"
misc_feature      270544..276847
/note="assembly_fragment"
misc_feature      276848..283151
/note="assembly_fragment"
misc_feature      283152..289455
/note="assembly_fragment"
misc_feature      289456..295759
/note="assembly_fragment"
misc_feature      295760..302063
/note="assembly_fragment"
misc_feature      302064..308367
/note="assembly_fragment"
misc_feature      308368..314671
/note="assembly_fragment"
misc_feature      314672..320975
/note="assembly_fragment"
misc_feature      320976..327279
/note="assembly_fragment"
misc_feature      327280..333583
/note="assembly_fragment"
misc_feature      333584..339887
/note="assembly_fragment"
misc_feature      339888..346191
/note="assembly_fragment"
misc_feature      346192..352495
/note="assembly_fragment"
misc_feature      352496..358799
/note="assembly_fragment"
misc_feature      358800..365103
/note="assembly_fragment"
misc_feature      365104..371407
/note="assembly_fragment"
misc_feature      371408..377711
/note="assembly_fragment"
misc_feature      377712..384015
/note="assembly_fragment"
misc_feature      384016..390319
/note="assembly_fragment"
misc_feature      390320..396623
/note="assembly_fragment"
misc_feature      396624..402927
/note="assembly_fragment"
misc_feature      402928..409231
/note="assembly_fragment"
misc_feature      409232..415535
/note="assembly_fragment"
misc_feature      415536..421839
/note="assembly_fragment"
misc_feature      421840..428143
/note="assembly_fragment"
misc_feature      428144..434447
/note="assembly_fragment"
misc_feature      434448..440751
/note="assembly_fragment"
misc_feature      440752..447055
/note="assembly_fragment"
misc_feature      447056..453359
/note="assembly_fragment"
misc_feature      453360..459663
/note="assembly_fragment"
misc_feature      459664..465967
/note="assembly_fragment"
misc_feature      465968..472271
/note="assembly_fragment"
misc_feature      472272..478575
/note="assembly_fragment"
misc_feature      478576..484879
/note="assembly_fragment"
misc_feature      484880..491183
/note="assembly_fragment"
misc_feature      491184..497487
/note="assembly_fragment"
misc_feature      497488..503791
/note="assembly_fragment"
misc_feature      503792..510095
/note="assembly_fragment"
misc_feature      510096..516399
/note="assembly_fragment"
misc_feature      516400..522703
/note="assembly_fragment"
misc_feature      522704..529007
/note="assembly_fragment"
misc_feature      529008..535311
/note="assembly_fragment"
misc_feature      535312..541615
/note="assembly_fragment"
misc_feature      541616..547919
/note="assembly_fragment"
misc_feature      547920..554223
/note="assembly_fragment"
misc_feature      554224..560527
/note="assembly_fragment"
misc_feature      560528..566831
/note="assembly_fragment"
misc_feature      566832..573135
/note="assembly_fragment"
misc_feature      573136..579439
/note="assembly_fragment"
misc_feature      579440..585743
/note="assembly_fragment"
misc_feature      585744..592047
/note="assembly_fragment"
misc_feature      592048..598351
/note="assembly_fragment"
misc_feature      598352..604655
/note="assembly_fragment"
misc_feature      604656..610959
/note="assembly_fragment"
misc_feature      610960..617263
/note="assembly_fragment"
misc_feature      617264..623567
/note="assembly_fragment"
misc_feature      623568..629871
/note="assembly_fragment"
misc_feature      629872..636175
/note="assembly_fragment"
misc_feature      636176..642479
/note="assembly_fragment"
misc_feature      642480..648783
/note="assembly_fragment"
misc_feature      648784..655087
/note="assembly_fragment"
misc_feature      655088..661391
/note="assembly_fragment"
misc_feature      661392..667695
/note="assembly_fragment"
misc_feature      667696..673999
/note="assembly_fragment"
misc_feature      674000..680303
/note="assembly_fragment"
misc_feature      680304..686607
/note="assembly_fragment"
misc_feature      686608..692911
/note="assembly_fragment"
misc_feature      692912..699215
/note="assembly_fragment"
misc_feature      699216..705519
/note="assembly_fragment"
misc_feature      705520..711823
/note="assembly_fragment"
misc_feature      711824..718127
/note="assembly_fragment"
misc_feature      718128..724431
/note="assembly_fragment"
misc_feature      724432..730735
/note="assembly_fragment"
misc_feature      730736..737039
/note="assembly_fragment"
misc_feature      737040..743343
/note="assembly_fragment"
misc_feature      743344..749647
/note="assembly_fragment"
misc_feature      749648..755951
/note="assembly_fragment"
misc_feature      755952..762255
/note="assembly_fragment"
misc_feature      762256..768559
/note="assembly_fragment"
misc_feature      768560..774863
/note="assembly_fragment"
misc_feature      774864..781167
/note="assembly_fragment"
misc_feature      781168..787471
/note="assembly_fragment"
misc_feature      787472..793775
/note="assembly_fragment"
misc_feature      793776..800079
/note="assembly_fragment"
misc_feature      800080..806383
/note="assembly_fragment"
misc_feature      806384..812687
/note="assembly_fragment"
misc_feature      812688..818991
/note="assembly_fragment"
misc_feature      818992..825295
/note="assembly_fragment"
misc_feature      825296..831599
/note="assembly_fragment"
misc_feature      831600..837903
/note="assembly_fragment"
misc_feature      837904..844207
/note="assembly_fragment"
misc_feature      844208..850511
/note="assembly_fragment"
misc_feature      850512..856815
/note="assembly_fragment"
misc_feature      856816..863119
/note="assembly_fragment"
misc_feature      863120..869423
/note="assembly_fragment"
misc_feature      869424..875727
/note="assembly_fragment"
misc_feature      875728..882031
/note="assembly_fragment"
misc_feature      882032..888335
/note="assembly_fragment"
misc_feature      888336..894639
/note="assembly_fragment"
misc_feature      894640..900943
/note="assembly_fragment"
misc_feature      900944..907247
/note="assembly_fragment"
misc_feature      907248..913551
/note="assembly_fragment"
misc_feature      913552..919855
/note="assembly_fragment"
misc_feature      919856..926159
/note="assembly_fragment"
misc_feature      926160..932463
/note="assembly_fragment"
misc_feature      932464..938767
/note="assembly_fragment"
misc_feature      938768..945071
/note="assembly_fragment"
misc_feature      945072..951375
/note="assembly_fragment"
misc_feature      951376..957679
/note="assembly_fragment"
misc_feature      957680..963983
/note="assembly_fragment"
misc_feature      963984..970287
/note="assembly_fragment"
misc_feature      970288..976591
/note="assembly_fragment"
misc_feature      976592..982895
/note="assembly_fragment"
misc_feature      982896..989199
/note="assembly_fragment"
misc_feature      989200..995503
/note="assembly_fragment"
misc_feature      995504..1001807
/note="assembly_fragment"
misc_feature      1001808..1008111
/note="assembly_fragment"
misc_feature      1008112..1014415
/note="assembly_fragment"
misc_feature      1014416..1020719
/note="assembly_fragment"
misc_feature      1020720..1027023
/note="assembly_fragment"
misc_feature      1027024..1033327
/note="assembly_fragment"
misc_feature      1033328..1039631
/note="assembly_fragment"
misc_feature      1039632..1045935
/note="assembly_fragment"
misc_feature      1045936..1052239
/note="assembly_fragment"
misc_feature      1052240..1058543
/note="assembly_fragment"
misc_feature      1058544..1064847
/note="assembly_fragment"
misc_feature      1064848..1071151
/note="assembly_fragment"
misc_feature      1071152..1077455
/note="assembly_fragment"
misc_feature      1077456..1083759
/note="assembly_fragment"
misc_feature      1083760..1090063
/note="assembly_fragment"
misc_feature      1090064..1096367
/note="assembly_fragment"
misc_feature      1096368..1102671
/note="assembly_fragment"
misc_feature      1102672..1108975
/note="assembly_fragment"
misc_feature      1108976..1115279
/note="assembly_fragment"
misc_feature      1115280..1121583
/note="assembly_fragment"
misc_feature      1121584..1127887
/note="assembly_fragment"
misc_feature      1127888..1134191
/note="assembly_fragment"
misc_feature      1134192..1140495
/note="assembly_fragment"
misc_feature      1140496..1146799
/note="assembly_fragment"
misc
```

ORGANISM	Homo sapiens		
REFERENCE	Enkayrtozoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulhelfia; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 142558)		
JOURNAL	Wiley		
COMMENT	Project: Submission Submitted (17-Jul-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hinxton@sanger.ac.uk On Jul 19, 2001 this sequence version replaced gi:14529732. ----- Genome Center Center: Sanger Centre Center code: SC Web site: http://www.sanger.ac.uk Contact: hinxton@sanger.ac.uk ----- Project information Center project name: d88b1p19 ----- Summary statistics Assembly program: XGAP; version 4.5 Sequencing vector: pLasmid; L08752; 100% of reads Chemistry: Dye-terminator; Big Dye; 99% of reads Insert size: 1378/2; 3.8% error; agarose-fp Consensus quality: 1392/5 bases at least Q40 Consensus quality: 140700 bases at least Q30 Consensus quality: 141500 bases at least Q20 Insert size: 1378/2; 3.8% error; agarose-fp Insert size: 1378/2; 3.8% error; agarose-fp Quality coverage: 7.73x in Q20 bases; sum-of-conlgs quality coverage: 5.52x in Q20 bases; agarose-fp ----- NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. ----- 1 2667: contig of 2667 bp in length * 2668 2767: gap of 100 bp * 2768 5694: contig of 2927 bp in length * 5695 5794: gap of 100 bp * 5795 7987: contig of 2193 bp in length * 7988 13350: gap of 5355 bp in length * 13350 136003: contig of 127816 bp in length * 136004 139240: contig of 3237 bp in length * 139241 139340: gap of 100 bp * 139341 142558: contig of 3218 bp in length. ----- Location/Qualifiers 1..142558 "Homo sapiens" /genbank="U000000000" /chromosome="11" /map="q42.11-q2.3" /clone="RP5-881P19" /clone_1lb="RPIC-5" 1..2667 /note="assembly-fragment:02382" 2768..5694 /note="assembly-fragment:02427" 5795..7987 /note="assembly-fragment:02438" 8088..135503 /note="assembly-fragment:02459" 136004..139240 /note="assembly-fragment:02490" 139241..142558 /note="assembly-fragment:02497" ----- BASE COUNT 33414 a 39141 c 35025 g 33673 t 505 others ORIGIN		


```

ACCESSION      Xq13.3-21.1 complete sequence.
VERSION        AL357115
KEYWORDS       AL357115.24  GI:126662359
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS        Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
TITLE          1 (bases 1 to 164012)
JOURNAL
COMMENT
Direct Submission
Submitted (02-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
UK.
Requester(s): AL357115, AL357115.24, AL357115.25, AL357115.26, AL357115.27, AL357115.28, AL357115.29, AL357115.30, AL357115.31, AL357115.32, AL357115.33, AL357115.34, AL357115.35, AL357115.36, AL357115.37, AL357115.38, AL357115.39, AL357115.40, AL357115.41, AL357115.42, AL357115.43, AL357115.44, AL357115.45, AL357115.46, AL357115.47, AL357115.48, AL357115.49, AL357115.50, AL357115.51, AL357115.52, AL357115.53, AL357115.54, AL357115.55, AL357115.56, AL357115.57, AL357115.58, AL357115.59, AL357115.60, AL357115.61, AL357115.62, AL357115.63, AL357115.64, AL357115.65, AL357115.66, AL357115.67, AL357115.68, AL357115.69, AL357115.70, AL357115.71, AL357115.72, AL357115.73, AL357115.74, AL357115.75, AL357115.76, AL357115.77, AL357115.78, AL357115.79, AL357115.80, AL357115.81, AL357115.82, AL357115.83, AL357115.84, AL357115.85, AL357115.86, AL357115.87, AL357115.88, AL357115.89, AL357115.90, AL357115.91, AL357115.92, AL357115.93, AL357115.94, AL357115.95, AL357115.96, AL357115.97, AL357115.98, AL357115.99, AL357115.100, AL357115.101, AL357115.102, AL357115.103, AL357115.104, AL357115.105, AL357115.106, AL357115.107, AL357115.108, AL357115.109, AL357115.110, AL357115.111, AL357115.112, AL357115.113, AL357115.114, AL357115.115, AL357115.116, AL357115.117, AL357115.118, AL357115.119, AL357115.120, AL357115.121, AL357115.122, AL357115.123, AL357115.124, AL357115.125, AL357115.126, AL357115.127, AL357115.128, AL357115.129, AL357115.130, AL357115.131, AL357115.132, AL357115.133, AL357115.134, AL357115.135, AL357115.136, AL357115.137, AL357115.138, AL357115.139, AL357115.140, AL357115.141, AL357115.142, AL357115.143, AL357115.144, AL357115.145, AL357115.146, AL357115.147, AL357115.148, AL357115.149, AL357115.150, AL357115.151, AL357115.152, AL357115.153, AL357115.154, AL357115.155, AL357115.156, AL357115.157, AL357115.158, AL357115.159, AL357115.160, AL357115.161, AL357115.162, AL357115.163, AL357115.164, AL357115.165, AL357115.166, AL357115.167, AL357115.168, AL357115.169, AL357115.170, AL357115.171, AL357115.172, AL357115.173, AL357115.174, AL357115.175, AL357115.176, AL357115.177, AL357115.178, AL357115.179, AL357115.180, AL357115.181, AL357115.182, AL357115.183, AL357115.184, AL357115.185, AL357115.186, AL357115.187, AL357115.188, AL357115.189, AL357115.190, AL357115.191, AL357115.192, AL357115.193, AL357115.194, AL357115.195, AL357115.196, AL357115.197, AL357115.198, AL357115.199, AL357115.200, AL357115.201, AL357115.202, AL357115.203, AL357115.204, AL357115.205, AL357115.206, AL357115.207, AL357115.208, AL357115.209, AL357115.210, AL357115.211, AL357115.212, AL357115.213, AL357115.214, AL357115.215, AL357115.216, AL357115.217, AL357115.218, AL357115.219, AL357115.220, AL357115.221, AL357115.222, AL357115.223, AL357115.224, AL357115.225, AL357115.226, AL357115.227, AL357115.228, AL357115.229, AL357115.230, AL357115.231, AL357115.232, AL357115.233, AL357115.234, AL357115.235, AL357115.236, AL357115.237, AL357115.238, AL357115.239, AL357115.240, AL357115.241, AL357115.242, AL357115.243, AL357115.244, AL357115.245, AL357115.246, AL357115.247, AL357115.248, AL357115.249, AL357115.250, AL357115.251, AL357115.252, AL357115.253, AL357115.254, AL357115.255, AL357115.256, AL357115.257, AL357115.258, AL357115.259, AL357115.260, AL357115.261, AL357115.262, AL357115.263, AL357115.264, AL357115.265, AL357115.266, AL357115.267, AL357115.268, AL357115.269, AL357115.270, AL357115.271, AL357115.272, AL357115.273, AL357115.274, AL357115.275, AL357115.276, AL357115.277, AL357115.278, AL357115.279, AL357115.280, AL357115.281, AL357115.282, AL357115.283, AL357115.284, AL357115.285, AL357115.286, AL357115.287, AL357115.288, AL357115.289, AL357115.290, AL357115.291, AL357115.292, AL357115.293, AL357115.294, AL357115.295, AL357115.296, AL357115.297, AL357115.298, AL357115.299, AL357115.300, AL357115.301, AL357115.302, AL357115.303, AL357115.304, AL357115.305, AL357115.306, AL357115.307, AL357115.308, AL357115.309, AL357115.310, AL357115.311, AL357115.312, AL357115.313, AL357115.314, AL357115.315, AL357115.316, AL357115.317, AL357115.318, AL357115.319, AL357115.320, AL357115.321, AL357115.322, AL357115.323, AL357115.324, AL357115.325, AL357115.326, AL357115.327, AL357115.328, AL357115.329, AL357115.330, AL357115.331, AL357115.332, AL357115.333, AL357115.334, AL357115.335, AL357115.336, AL357115.337, AL357115.338, AL357115.339, AL357115.340, AL357115.341, AL357115.342, AL357115.343, AL357115.344, AL357115.345, AL357115.346, AL357115.347, AL357115.348, AL357115.349, AL357115.350, AL357115.351, AL357115.352, AL357115.353, AL357115.354, AL35711
```

ACCESSION	AL512502
VERSION	AL512502.4
KEYWORDS	HTG; HTGS; PASEL; HTGS; CANCELLED.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Bernard, M., Kellum, J., Chordato, C., Catalina, V., Verheggen, B., Euseles Mammals I, Eutheria II, Primates I, Carnivora I, Homiidae, Homo I (bases 1 to 177263)
TITLE	Health.P.
JOURNAL	Direct Submission Submitted (09-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire CB10 1SA, UK. E-mail enquiries: humm@sanger.ac.uk Requests: ffr@sanger.ac.uk Genome version replaced g1:13092346. genome Center
COMMENT	Center: Sanger Centre Web site: http://www.sanger.ac.uk Contact: humm@sanger.ac.uk Project Information center project name: BA00785 Assembly program: XMAP4, version 4.5 Sequencing vector: plasmid: L08752; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Consensus quality: 176599 bases at least Q40 Consensus quality: 176106 bases at least Q30 Consensus quality: 176347 bases at least Q20 Insert size: 176703; sum-of-contigs sum-of-contigs: 176703; sum-of-contigs-fp Quality coverage: 7.03 in Q20 bases; sum-of-contigs Quality coverage: 7.16x in Q20 bases; avgase -fp
FEATURES	* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Contig numbers are assigned sequentially from contig 1, but the exact sizes of the gaps between these records will be updated with the finished sequence, as soon as it is available and the accession number will be preserved. 1 8339: contig of 8339 bp in length 2 100 bp 3 8340 8439: gap of 100 bp 4 8440 6916040: contig of 606601 bp in length 5 69161 6916040: contig of 606601 bp in length 6 69141 108593: contig of 39453 bp in length 7 108594 108693: gap of 100 bp 8 108694 113120: contig of 4427 bp in length 9 113121 113220: gap of 100 bp 10 113221 148811: contig of 35591 bp in length 11 148812 148911: gap of 100 bp 12 148912 177263: contig of 28352 bp in length. 1 localQualifiers 1 localQualifiers 1 /organism='Homo sapiens' 2 /db_xref='taxon:9606' 3 /chromosome='X' 4 /clone='RP11-807B6' 5 /clone_id='RP11-807B6' 6 /feature='assembly:fragment:01802' 7 /feature:chain:1' 8 /note='assembly:fragment:02140' 9 /feature:chain:1' 10 69141..108593 11 /note='assembly:fragment:03321' 12 108594..113120 13 /note='assembly:fragment:00292' 14 /feature:chain:1' 15 113221..148811 16 /note='assembly:fragment:01412'

```

fragment_chain:1*
misc_feature      168912..177263 fragment:03043
                     /gene="assembly1"
                     /locus_tag="R23-449P12"
BASE COUNT      51647 a 31389 c 34060 g 59664 t      503 others
ORIGIN
Alignment_scores:
    quality:      54.00      length:      11
    ratio:        9.909      gaps:          0
Percent similarity: 90.909 Percent identity: 72.727
Alignment_block:
US-09-471-276-831_COPY_1_16 x AL1512502 ..
Align seg 1/1 to: AL1512502 from: 1 to: 177263
6 ValPheLeuLeuTrpGlyValTrpGly 16
|||||:|||||:|||||:|||||:|||||:
4897 GTTACCTTTCTTGGCGGATTTGTCGGCG 4929

seq_name: gb_hlg:AC079557
seq_documentation_block:
AC079557.1 165839 bp DNA HTG 02-SEP-2000
DEFINITION uncloned plasmid.
ACCESSION AC079557
VERSION AC079557.1 GI:9664922
KEYWORDS HTG: HTGS, PHASE1: HTGS, DNAPT.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 165839)
AUTHORS DOE Joint Genome Institute.
JOURNAL Unpublished
TITLES Sequencing of Mouse
REFERENCE 2 (bases 1 to 185839)
AUTHORS DOE Joint Genome Institute.
JOURNAL Unpublished
TITLES Sequencing of Mouse
SUBMITTER 02-SEP-2000 Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
Project Information
Center Project Name: 1909420
Center clone name: RNCr23_449P12
-----
Summary Statistics
Consensus quality: 172503 bases at least Q40
Consensus quality: 17766 bases at least Q20
Consensus quality: 179073 bases at least Q20
Consensus quality: 184139 sum-of-contigs estimation
Estimated insert size: 184139
Quality coverage: 7.81 in Q20 bases; average-tp estimation
NOTE: This is a 'working draft' sequence. It currently
consists of 16 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
dashes. The gaps are not necessarily in the same order
as they appear in the sequence.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 1081: contig of 1081 bp in length
* 1082 1181: gap of unknown length
* 1182 4298: contig of 3117 bp in length
* 4299 4398: gap of unknown length

```

```

* 4399 6732: contig of 2334 bp in length
* 6733 6832: gap of unknown length
* 6833 11626: contig of 4794 bp in length
* 11626 11727: contig of 101 bp in length
* 11727 15445: contig of 3819 bp in length
* 15445 15546: gap of unknown length
* 15546 20007: contig of 4362 bp in length
* 20007 20108: gap of unknown length
* 20108 26156: contig of 6049 bp in length
* 26156 33058: gap of unknown length
* 33058 33059: contig of 6702 bp in length
* 33059 33060: gap of unknown length
* 33060 42014: contig of 8956 bp in length
* 42014 42115: gap of unknown length
* 42115 50688: contig of 8574 bp in length
* 50688 50788: gap of unknown length
* 50788 61289: contig of 10510 bp in length
* 61289 61388: gap of unknown length
* 61388 77220: contig of 15871 bp in length
* 77220 77221: gap of unknown length
* 77221 95693: contig of 18324 bp in length
* 95693 95793: gap of unknown length
* 95793 119999: contig of 24206 bp in length
* 120000 120099: gap of unknown length
* 120099 150273: contig of 30174 bp in length
* 150273 150373: gap of unknown length
* 150373 150374: contig of 35466 bp in length.
Location: 165839
Features:
source
1..165839
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="R23-449P12"
/clone_11b="R23-449P12"
BASE COUNT 48719 a 44398 c 42466 g 49655 t 1501 others
ORIGIN
Alignment_scores:
    quality:      54.00      length:      12
    ratio:        9.909      gaps:          0
Percent similarity: 91.667 Percent identity: 66.667
Alignment_block:
US-09-471-276-831_COPY_1_16 x AC079557 ..
Align seg 1/1 to: AC079557 from: 1 to: 185839
5 ValPheLeuLeuLeuTrpGlyValTrpGly 16
|||||:|||||:|||||:|||||:|||||:
151627 CTCATTTCTTCTTGGCGGATTTGTCGGCG 151627
seq_name: gb_pr:AL136307
seq_documentation_block:
LOCUS AL136307 186882 bp DNA PRI 01-AUG-2000
DEFINITION Human DNA sequence from clone R3-3808B on chromosome 6p24.1-25.3
CONTAINS A gene encoding the protein Neurtelin, which is involved in
protection of neurite outgrowth, a tyrosine kinase (PKX2) and
protection of neural mRNAs, 4 CpG islands, 5'UTR, 3'UTR and GSSs,
complete sequence.
ACCESSION AL136307
VERSION AL136307.12 GI:8574124
KEYWORDS HTG: CpG island; Neurtlin; PKX2.
SOURCE human.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 186882)
AUTHORS Kikawaga; Matsuzo; Chordella; Granata; Vertehrata; Euclescomi;
Brown, J.
TITLES Direct Submission
JOURNAL Submitted (28-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: hinxton@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

```

COMMENT

On Jun 20, 2000 this sequence version replaced gi:8573783. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission (e.g. the 1299 bp overlap between RP3-380B8 and RP3-380B8.3 only a small overlap as described above). This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' tag. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information from the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Project. The human chromosome 6 map can be found at <http://www.sanger.ac.uk/projects/chr6/> RP3-380B8 is from the library RP3-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pCYPC2

This sequence is the entire insert of clone RP3-380B8. The true left end of clone RP3-380B8 is at 14150. In this sequence, The true right end of clone RP3-380B8 is at 57138. In this sequence.

FEATURES

source

1..166882
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="P283-380B8.3"
/clone="RP3-380B8"
/clone_id="RP3-380B8.3"
247..465
/note="MIR repeat: matches 16..236 of consensus"
MIR repeat: matches 16..236 of consensus"
1442..2391
/note="L1M83 repeat: matches 5281..6174 of consensus"
L1M83 repeat: matches 5281..6174 of consensus"
2435..2499
/note="MIR repeat: matches 77..145 of consensus"
MIR repeat: matches 77..145 of consensus"
2540..3786
/note="L1P42 repeat: matches 4900..6146 of consensus"
L1P42 repeat: matches 4900..6146 of consensus"
2542..3786
/note="L1P45 repeat: matches 4900..6143 of consensus"
L1P45 repeat: matches 4900..6143 of consensus"
4034..4639
/note="L1M3 repeat: matches 4022..4632 of consensus"
L1M3 repeat: matches 4022..4632 of consensus"
4654..4953
/note="AluX repeat: matches 1..300 of consensus"
AluX repeat: matches 1..300 of consensus"
4954..6436
/note="L1M49 repeat: matches 4632..6308 of consensus"
L1M49 repeat: matches 4632..6308 of consensus"
6522..6820
/note="MUTD repeat: matches 14..283 of consensus"
MUTD repeat: matches 14..283 of consensus"
6821..6990
/note="MER67A repeat: matches 374..543 of consensus"
MER67A repeat: matches 374..543 of consensus"
6991..7290
/note="MUTD repeat: matches 308..505 of consensus"
MUTD repeat: matches 308..505 of consensus"
7290..7462
/note="MIR repeat: matches 71..262 of consensus"
MIR repeat: matches 71..262 of consensus"
7357..7462
/note="MIR repeat: matches 71..177 of consensus"
MIR repeat: matches 71..177 of consensus"
8007..8093

1/mote="L1M4 repeat: matches 5688..5772 of consensus"
L1M4 repeat: matches 5688..5772 of consensus"
8094..8239
/note="L1M8 repeat: matches 6029..6173 of consensus"
L1M8 repeat: matches 6029..6173 of consensus"
8239..8996
/note="L2 repeat: matches 2637..2699 of consensus"
L2 repeat: matches 2637..2699 of consensus"
8996..9199
/note="MER33 repeat: matches -9..192 of consensus"
MER33 repeat: matches -9..192 of consensus"
9200..9559
/note="MER1B repeat: matches 1..364 of consensus"
MER1B repeat: matches 1..364 of consensus"
9560..10364
/note="MER33 repeat: matches 192..324 of consensus"
MER33 repeat: matches 192..324 of consensus"
10364..10771
/note="MUTG repeat: matches 65..512 of consensus"
MUTG repeat: matches 65..512 of consensus"
10771..10875
/note="MIR repeat: matches 65..512 of consensus"
MIR repeat: matches 65..512 of consensus"
10875..10875
/note="L1M40 repeat: matches 5479..6315 of consensus"
L1M40 repeat: matches 5479..6315 of consensus"
12979..13363
/note="AluX repeat: matches 9..290 of consensus"
AluX repeat: matches 9..290 of consensus"
13363..13787
/note="MIR repeat: matches 15..262 of consensus"
MIR repeat: matches 15..262 of consensus"
14035..14188
/note="MER3A repeat: matches 43..189 of consensus"
MER3A repeat: matches 43..189 of consensus"
14785..14859
/note="L1M53 repeat: matches 5785..5859 of consensus"
L1M53 repeat: matches 5785..5859 of consensus"
15050..15310
/note="L1M33 repeat: matches 5525..5777 of consensus"
L1M33 repeat: matches 5525..5777 of consensus"
15317..15774
/note="L1M4 repeat: matches 3951..4454 of consensus"
L1M4 repeat: matches 3951..4454 of consensus"
15317..15774
/note="L1 repeat: matches 3951..4433 of consensus"
L1 repeat: matches 3951..4433 of consensus"
15774..16616
/note="L2 repeat: matches 2664..2750 of consensus"
L2 repeat: matches 2664..2750 of consensus"
16616..17138
/note="match: GSS: Em:A0541664"
match: GSS: Em:A0541664"
17138..17193
/note="MIR repeat: matches 78..133 of consensus"
MIR repeat: matches 78..133 of consensus"
17193..18904
/note="L1M3 repeat: matches 193..466 of consensus"
L1M3 repeat: matches 193..466 of consensus"
18904..19230
/note="MIR repeat: matches 1..310 of consensus"
MIR repeat: matches 1..310 of consensus"
19230..19419
/note="MUTG repeat: matches 1..193 of consensus"
MUTG repeat: matches 1..193 of consensus"
19419..20155
/note="L1M8 repeat: matches 6029..6173 of consensus"
L1M8 repeat: matches 6029..6173 of consensus"
20155..20184
/note="match: ESTs: Em:R79825"
match: ESTs: Em:R79825"
20184..20482
/note="L1M3 repeat: matches 193..466 of consensus"
L1M3 repeat: matches 193..466 of consensus"
20482..20806
/note="L1M3 repeat: matches 193..466 of consensus"
L1M3 repeat: matches 193..466 of consensus"
20806..20806
/note="L1M3 repeat: matches 193..466 of consensus"
L1M3 repeat: matches 193..466 of consensus"
20806..20806
/note="L1M3 repeat: matches 193..466 of consensus"
L1M3 repeat: matches 193..466 of consensus"


```

Center Project name: 16917
Center Clone name: 21_G_15
----- Summary Statistics -----
Sequencing vector: M13: M77815, 100% of reads
Sequencing vector: pTZ19: 16908, 100% of reads
Assembly: 13967 bases at least Q30
Consensus quality: 13967 bases at least Q30
Consensus quality: 140370 bases at least Q30
Consensus quality: 142846 bases at least Q30
Insert size: 193000; agarose-fp
Insert size: 144056; sum-of-coverage
Quality coverage: 4.2 in Q30 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* as soon as it is available and the accession number M11
* be preserved
1 15351: contig of 15351 bp in length
13532 15451: gap of 100 bp
13452 16828: contig of 1377 bp in length
16829 16928: gap of 100 bp
15929 16033: contig of 1011 bp in length
15930 16033: contig of 1011 bp in length
16040 16688: contig of 1649 bp in length
16689 19788: gap of 100 bp
19789 21805: contig of 2017 bp in length
21806 21905: gap of 100 bp
21906 24023: contig of 2088 bp in length
24024 24023: gap of 100 bp
24024 24023: contig of 1531 bp in length
27405 27404: gap of 100 bp
27405 32590: contig of 5186 bp in length
32591 32690: gap of 100 bp
32691 37217: contig of 4527 bp in length
37218 37317: gap of 100 bp
37318 43108: contig of 5791 bp in length
43109 43209: contig of 4455 bp in length
43209 47763: contig of 4555 bp in length
47764 47863: gap of 100 bp
47864 53820: contig of 5957 bp in length
53821 53920: gap of 100 bp
53921 58914: contig of 4994 bp in length
58915 59014: gap of 100 bp
59015 64835: contig of 2634 bp in length
64836 64835: contig of 9533 bp in length
64836 94988: contig of 9533 bp in length
94989 95088: gap of 100 bp
95089 109604: contig of 14516 bp in length
109605 109704: gap of 100 bp
109705 122448: contig of 12724 bp in length
122449 122528: gap of 100 bp
122529 1303527: contig of 16999 bp in length
130353 130352: gap of 100 bp
130628 145856: contig of 6229 bp in length.
Location/Qualifiers
1. 145856
/organism: "Homo sapiens"
/db_xref: "taxon:9606"
/chromosome: "18"
/amp: "18"
/clone: "RP11-21G15"
/clone_1lb: "RP11-11 Human Whole BAC"
1. 15351
/clone_1lb: "RP11-11 Human Whole BAC"
/clone_end: "sp"
vector_side: "left"
15452..16828
/note: "assembly_fragment"

```

```

misc_feature 16929..17939
/note: "assembly_fragment"
misc_feature 18040..19688
/note: "assembly_fragment"
misc_feature 19789..21805
/note: "assembly_fragment"
misc_feature 21906..23993
/note: "assembly_fragment"
misc_feature 24094..27304
/note: "assembly_fragment"
misc_feature 27405..32590
/note: "assembly_fragment"
misc_feature 32691..37217
/note: "assembly_fragment"
misc_feature 37318..43108
/note: "assembly_fragment"
misc_feature 43209..47763
/note: "assembly_fragment"
misc_feature 47864..53820
/note: "assembly_fragment"
misc_feature 53921..58914
/note: "assembly_fragment"
misc_feature 59015..64835
/note: "assembly_fragment"
misc_feature 64836..94988
/note: "assembly_fragment"
misc_feature 95089..109604
/note: "assembly_fragment"
misc_feature 109705..122528
/note: "assembly_fragment"
misc_feature 122529..139527
/note: "assembly_fragment"
misc_feature 139628..145856
/note: "assembly_fragment"
misc_feature 145857..17939
/note: "assembly_fragment"
BASE COUNT 41606 a 30265 c 28947 g 4333 t 1805 others
ORIGIN
alignment_scores:
Quantity: 53.00 Length: 15
Percent Similarity: 73.33 Percent Identity: 60.000
alignment_block:
US-09-471-276-831_COPY_1.16 x AC073948 ..
Align seg 1/1 to: AC073948 from: 1 to: 145856
2 SeqName: US09471276831COPY116
|||||
|||||
14470 TCATCTGTATATCTTCTTCTTACCTTGCAACACGATGGGGT 1514
seq_name: gb_hhg:AC019241
seq_documentation_block:
LOCUS AC019241 155804 bp DNA HTG 10-NOV-2000
DEFINITION Homo sapiens clone RP11-45120, WORKING DRAFT SEQUENCE, 6 unordered
regions.
ACCESSION AC019241
VERSION AC019241.5 GI:11136802
KEYWORDS HTG: HTGS_PHASET; HTGS_DRAFT.
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa;
Mammalia; Primates; Hominidae; Homininae; Homininae; Homo.
REFERENCE
1 Birren, B., Linton, L., Nusbaum, C., Lander, E.,
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 155804)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,

```



```

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165277 bases at least Q40
Consensus quality: 165932 bases at least Q40
Consensus quality: 167273 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 167808; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; agarose-fp
Quality coverage: 5.7 in Q20 bp.
NOTE: This is a 'working draft' sequence. It currently
consists of 5 contigs. The true order of the pieces
is not known and these contigs are arbitrary.
Arbitrary gaps between contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
• 85800 85899; contig of 85799 bp in length
• 85900 85900; gap of 100 bp
• 85901 85901; contig of 100 bp in length
• 96515 96515; gap of 100 bp
• 96515 96516; contig of 19602 bp in length
• 116317 116416; gap of 100 bp
• 116317 116416; gap of 100 bp
• 116417 141745; contig of 25129 bp in length
• 141746 141845; gap of 100 bp
• 141846 168208; contig of 26363 bp in length.
Loc:1030Qualifiers
FEATURES
source
/organism:"Homo sapiens"
/db_xref:"taxon:9606"
/chromosome:"8"
/map:"8"
/clone:"RP11-272p15"
/contig:"110" RPT-11 Human Male BAC"
misc_feature
/note:"assembly_fragment"
clone_end:spg
vector_side:left"
misc_feature
85900..96514
/note:"assembly_fragment"
misc_feature
96715..116316
/note:"assembly_fragment"
misc_feature
116417..141745
/note:"assembly_fragment"
141846..168208
misc_feature
/note:"assembly_fragment"
clone_end:17
vector_side:right"
BASE COUNT
52087 a 51238 c 30497 g 53965 t 401 others
ORIGIN
alignment_scores:
Quality: 53.00 Length: 10
Ratio: 5.300 Gaps: 0
Percent identity: 100.000 Percent identity: 70.000
alignment_block:
US-09-471-276-831_COPY_1_16 x AC024359
Align seq 1/1 to: AC024359 from: 1 to: 168208
6 ValPhdCLeuLeuAurTrrpQlyValThrTrrp 15
142177 CTTTTCACATTCATCTGCGACATACATGCG 142206
seq_name: gb_hcg:AC023569
seq_documentation_block:
LOCUS AC023569 169607 bp DNA HTG 06-APR-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-513M1 map 18, WORKING DRAFT
SEQUENCE 17 unordered pieces.
AC023569

```

VERSION
AC023569.2 GI:7466913
HTG: HTGS_DRAFT.
SOURCE
human.

ORGANISM

Homo sapiens
Mammalia: Primates; Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates, Catarrhini, Hominoidea, Homo.

REFERENCE

1 (bases 1 to 169607)
Birren,B., Linton,L., Nussbaum,C. and Lander,E.

TITLES

JOURNAL

ARTWORKS

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES


```

      5210      8324: contig of 3115 bp in length
      8325 8424: gap of 100 bp
      11015 11014: contig of 2590 bp in length
      11015 11114: gap of 100 bp
      11115 11335: contig of 220 bp in length
      13395 13494: gap of 100 bp in length
      13495 17701: contig of 4207 bp in length
      17702 17801: gap of 100 bp in length
      17802 22017: contig of 4216 bp in length
      22018 22117: gap of 100 bp
      22118 25805: contig of 3686 bp in length
      25806 25905: gap of 100 bp
      25906 34369: gap of 100 bp in length
      34370 34369: gap of 100 bp in length
      44783 44882: gap of 100 bp in length
      44883 58944: contig of 14062 bp in length
      58945 59044: gap of 100 bp
      59045 73155: contig of 14115 bp in length
      73156 73280: gap of 100 bp in length
      73281 88114: gap of 100 bp in length
      88115 103661: contig of 15547 bp in length
      103662 103761: gap of 100 bp in length
      103762 134075: contig of 30314 bp in length
      134076 134175: gap of 100 bp
      134176 165607: contig of 35432 bp in length.

FEATURES
SOURCE
      134176..165607: contig of 35432 bp in length.
      Location/Qualifiers
      +organelle="Homo sapiens"
      +db_xref="taxon:9606"
      +chromosome="18"
      +map="18"
      +clone="RP11-513M1"
      +clone_lib="RP11-11 Human Male BAC"
      1..1050
      +note="assembly_fragment"
      1131..2617
      +note="assembly_fragment"
      2718..5109
      +note="assembly_fragment"
      5210..8324
      +note="assembly_fragment"
      8425..11014
      +note="assembly_fragment"
      11115..11334
      +note="assembly_fragment"
      13495..17701
      +note="assembly_fragment"
      17802..22017
      +note="assembly_fragment"
      22118..25805
      +note="assembly_fragment"
      25906..34369
      +note="assembly_fragment"
      34370..44782
      +note="assembly_fragment"
      44883..58944
      +note="assembly_fragment"
      59045..73155
      +note="assembly_fragment"
      73260..88014
      +note="assembly_fragment"
      88115..103661
      +note="assembly_fragment"
      103762..134075
      +note="assembly_fragment"
      134176..165607
      +note="assembly_fragment"

BAGE COUNT      50874 a 32780 c 33552 g 50796 t 1605 others
ORIGIN
alignment_scores:
      Quality: 53.00      Length: 15
      Ratio: 4.818      Gaps: 0
      Percent Similarity: 73.33      Percent Identity: 60.000

alignment_block:
US-05-471-276-831_COPY_1_16 x AC033569
Align seg 1/1 to: AC032569 from: 1 to: 165607
2 SerMetLeuValAlaPheLeuLeuLeuTrpGlyValAlaThrTrpGly 16
||||||| ::::::::::::::::::::| |||||
165938 TCATCCTCATATCTCTCTCTCTGTTGGAGACAGCGTGGCGT 165642

seq_name: gB_Htg:AC092580
seq_documentation_block:
LOCUS      AC092580      172005 bp      DNA
DEFINITION      Homo sapiens chromosome 2 clone RP11-16024, WORKING DRAFT SEQUENCE.
ACCESSION      AC092580
VERSION      AC092580.1 GI:14916165
KEYWORDS      HTG: HTS_PNASeq1: HTS: DRAFT.
SOURCE      human
ORGANISM      human
      Homo sapiens
      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
      1 (bases 1 to 172005)
      The sequence of Homo sapiens clone
      2 (bases 1 to 172005)
      Unpublished
      Reference:
      Title:
      Journal:
      Comment:
      Submitted (19-JUL-2001) Genome Sequencing Center, Washington
      University School of Medicine, 4444 Forest Park Parkway, St. Louis,
      MO 63108, USA
      On Jul 19, 2001 this sequence version replaced gi:7107944.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Contact: Dr. James Watson, jwatson@wustl.edu/gsc/index.shtml
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: R.NH0016024
Drafting center: MIBR
----- Summary Statistics -----
Sequencing vector: M13: 57%
Sequencing strategy: Shotgun
Chemistry: Dye-terminator Big Dye: 98% of reads
Assembly program: Phrap: Version 0.990319
Consensus quality: 171007 bases at least Q40
Consensus quality: 171565 bases at least Q40
Consensus quality: 171787 bases at least Q20
Insert size: 173580: agarose-fp
Insert size: 173580: agarose-fp
Insert size: 173580: agarose-fp
Quality coverage: 9.32 in Q20 bases: sum-of-contigs
Quality coverage: 9.66 in Q20 bases: sum-of-contigs
----- NOTE: This is a 'working draft' sequence. It currently
consists of 5 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. The gaps are marked with runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```



```

repeat_L_region /rpl_family="HERV1"
complement(4774..5479)
repeat_L_region /rpl_family="BTR5"
complement(2488..5766)
repeat_L_region /rpl_family="HERV1"
complement(5391..7000)
repeat_L_region /rpl_family="HERV1"
7360..8042
repeat_L_region /rpl_family="L2"
complement(8077..8416)
repeat_L_region /rpl_family="THEIA"
8538..8976
repeat_L_region /rpl_family="MERAD3"
complement(8977..9268)
repeat_L_region /rpl_family="ALUSg"
9269..9724
repeat_L_region /rpl_family="MERAD"
9725..9988
repeat_L_region /rpl_family="L2"
10200..10245
repeat_L_region /rpl_family="MLT2B1"
10674..11101
repeat_L_region /rpl_family="MER4E"
11012..11265
repeat_L_region /rpl_family="MLT2B1"
11086..11181
unSURE /note="30 qual SINGL region"
11126..11163
unSURE /note="11163 clone coverage"
11222..11345
unSURE /note="11345 clone coverage"
11262..11296
repeat_L_region /rpl_family="MLT2B1"
11762..11862
repeat_L_region /rpl_family="MLT2B2"
complement(12099..12225)
repeat_L_region /rpl_family="MLT2B2"
complement(12226..12515)
repeat_L_region /rpl_family="ALUS"
complement(12516..13835)
repeat_L_region /rpl_family="Char11c1a"
15559..15694
repeat_L_region /rpl_family="Cp-rich"
complement(15716..15771)
repeat_L_region /rpl_family="L1Pb1"
15772..16017
repeat_L_region /rpl_family="L1Pb1"
16018..16038
repeat_L_region /rpl_family="AT-rich"
complement(16358..16848)
repeat_L_region /rpl_family="L2"
16911..17192
repeat_L_region /rpl_family="GMAA)"
17119..17148
repeat_L_region /rpl_family="AT-rich"
complement(17644..18010)
repeat_L_region /rpl_family="L1M4"
complement(18045..18487)
repeat_L_region /rpl_family="L1M4a"
18488..18515
repeat_L_region /rpl_family="AT-rich"
18938..19306
repeat_L_region /rpl_family="MLT1K"
19582..19876
repeat_L_region /rpl_family="ALUS"
20037..20081
repeat_L_region /rpl_family="AT-rich"
complement(20153..20224)
repeat_L_region /rpl_family="L1M3"
complement(20394..20635)
repeat_L_region /rpl_family="L1M4"
complement(20717..20874)
repeat_L_region /rpl_family="L1M4"

```

```

repeat_L_region complement(20927..21384)
repeat_L_region /rpl_family="L1M3c"
22737..22790
repeat_L_region /rpl_family="AT-rich"
22875..22905
repeat_L_region /rpl_family="AT-rich"
23424..23501
repeat_L_region /rpl_family="MR"
24358..24385
repeat_L_region /rpl_family="AT-rich"
complement(24715..24880)
repeat_L_region /rpl_family="MR"
complement(25033..25141)
repeat_L_region /rpl_family="MR"
complement(25158..25311)
repeat_L_region /rpl_family="L1Pa13"
complement(25883..25978)
repeat_L_region /rpl_family="MR"
complement(26056..27011)
repeat_L_region /rpl_family="L1M4"
complement(28909..28989)
repeat_L_region /rpl_family="L1M3"
complement(30023..30225)
repeat_L_region /rpl_family="MR"
30337..30860
repeat_L_region /rpl_family="r19ger7"
30867..30954
repeat_L_region /rpl_family="r19ger7"
31873..31896

alignment_scores:
Quality: 53.00 length: 10
Ratio: 5.300 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-09-471-276-831_COPY_1_16 x AC016080 ..
Align seg 1/1 to: AC016080 from: 1 to: 176921

6 ValPheLeuLeuATrPGlyValThrTTP 15
:::|||||:|||||:|||||:|||||:
111951 CTTTTCATCTGCGGCAATTCATCG 111980

seq_name: gb_pl:ATCHRIV40
seq_documentation_block:
LOCUS ATCHRIV40 197405 bp DNA PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 40.
ACCESSION AL161540
VERSION AL161540.2 GI:7268224
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotys.
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 197405)
AUTHORS
EU Arabidopsis sequencing project.
TITLES
Direct Submission
JOURNAL
Submitted (2000) MGP, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail:
lmeckewitz.biochem.mpg.de; mayer@mpg.biochem.mpg.de
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK.
E-mail: michael.bevan@bsrc.ac.uk
COMMENT
Information on performance of analysis and a more detailed
description of the data and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.jic.ac.uk/chr3chr4chr5/
this fragment has an overlap with ATCHRIV39 at the 5' end and an
overlap with ATCHRIV41 at the 3' end.
FEATURES
Source
1..197405
Location/Qualifiers

```



```

repeat_region complement(34085..34439)
repeat_region /rpt_family="THEB1"
34573..34645
/rpt_family="THEB1"
repeat_region /rpt_family="THEB1"
36413..36768
repeat_region /rpt_family="THEB1-int"

Alignment_scores:
Quality: 53.00 Length: 10
Ratio: 5.300 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 70.000

alignment_block:
US-09-471-2276-831_COPY_1_Lib x NC012410/rev ..
Align seg 1/2 to reverse of: NC012410 from: 1 to: 220715

6 ValPhleLeuLeuTIPoLYvalTHrP 15
:::|||||:::|||||:::|||||:::|||||
87840 CTTTCCTACTGATCTGTGGCAATTACATCG 87841

seq_name: gb_in:CEI27E9

seq documentation_block:
LOCUS CEI27E9 22245 bp DNA INV 20-JUN-2001
DEFINITION Caenorhabditis elegans cosmid 727E9, complete sequence.
ACCESSION Z82059
VERSION Z82059.1 GI:1914550
KEYWORDS HTG.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE Rhabditiidae; Nematoda; Chromadorea; Rhabditiida;
Rhabditiidae; Rhabditiidae; Pelodictyinae; Caenorhabditis.
1 (bases 1 to 22245);
none.
genome sequence of the nematode C. elegans: A platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (1996), 2012-2018 (1998)
MHC2.elegans Sequencing Consortium.
2 (bases 1 to 22245)
Lloyd,C.R.
Direct Submission
Submitted (05-NOV-1996) Nematode Sequencing Project, Sanger Centre
Hinxton, Cambridgeshire CB10 1HQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jleas@sanger.ac.uk or rjemman@wustl.edu
On Mar 30, 1997, this sequence version replaced gi:1666020.
This submission has been replaced by a new submission, using
predictions from Genefinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye-terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
The following notes apply to this submission.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone 727E9 is at 19189 in
sequence AL032655.
The true right end of clone 727E9 is at 16950 in
sequence Z92813. This sequence (1..104) overlaps with the end of
sequence AL032655.
The end of this sequence (22142..22245) overlaps with the start of
sequence Z92813.
For a graphical representation of this sequence and its analysis
see:- http://wormbase.sanger.ac.uk/perl/acc/cilegans/scq/sequence?
name=727E9.

```

source 1. 22245
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="III"
/location="12769.1-12769.2"
/gene="C12769.2"
/note="predicted using GeneFinder
Similarly to bovine ubiquitinol-cytochrome c reductase
complex 11 KD protein (SM:P00126)
cDNA EST yk62802.5 comes from this gene"
/protein_id="CAB04873.1"
/db_xref="GI:3880275"
/db_xref="SPPRMS:045864"
/translat="SSSHSEPEADVYQITGYREKCDHYETFSILDECNRNRSK
SUTCECTGADVYVYHLDKADMPKASLK"
join(1296, 1335,1664, 1528,1574, 1794,1848, 2406)
/gene="C12769.1"
/note="predicted using GeneFinder
Similarly to bovine Ap/APP carrier protein (SM:P2307),
contains similarity to Pfam domain: P00153 (Mitochondrial
carrier protein), Score=382.1, E-value=1.8e-111, N=1
EMBL:216981 comes from this gene: cDNA EST yk61645
EMBL:486914 comes from this gene: cDNA EST EMBL:216445
cDNA EST yk935.5 comes from this gene: cDNA EST
EMBL:100728 comes from this gene: cDNA EST CEMS14F comes
from this gene
cDNA EST CEMS27F comes from this gene: cDNA EST yk7365.5
cDNA EST yk11964.3 comes from this gene: cDNA EST yk607.3
cDNA EST yk11964.3 comes from this gene: cDNA EST yk607.3
cDNA EST yk11964.3 comes from this gene: cDNA EST
EMBL:216981 comes from this gene: cDNA EST
EMBL:216981 comes from this gene: cDNA EST yk955.3 comes
from this gene
cDNA EST CEMS45R comes from this gene: cDNA EST CEMS577RB
cDNA EST CEMS45R comes from this gene: cDNA EST CEMS577RB
cDNA EST EMBL:100927 comes from this gene: cDNA EST
CEMS45R comes from this gene: cDNA EST EMBL:101509 comes
from this gene
cDNA EST CEMS46F comes from this gene: cDNA EST CEMS322PF
cDNA EST CEMS46F comes from this gene: cDNA EST CEMS322PF
cDNA EST CEMS59F comes from this gene: cDNA EST CEMS59F comes from this
gene
cDNA EST CEMS56F comes from this gene: cDNA EST yk6283.3
cDNA EST yk6719.3 comes from this gene: cDNA EST yk7565.3
cDNA EST yk6719.3 comes from this gene: cDNA EST yk7565.3
comes from this gene: cDNA EST yk7191.3 comes from this
gene
cDNA EST yk59610.3 comes from this gene: cDNA EST
yk10415.3 comes from this gene: cDNA EST yk60b7.5 comes
from this gene
cDNA EST yk6216.5 comes from this gene: cDNA EST yk6216.5
comes from this gene: cDNA EST yk64912.5 comes from this
gene
cDNA EST yk6761.5 comes from this gene: cDNA EST yk6769.5
comes from this gene: cDNA EST yk75d7.5 comes from this
gene
cDNA EST yk7265.5 comes from this gene: cDNA EST yk7567.5
comes from this gene: cDNA EST yk8279.5 comes from this
cDNA EST yk84d7.5 comes from this gene: cDNA EST yk8998.5
comes from this gene: cDNA EST yk90c1.5 comes from this
gene
cDNA EST yk96e10.5 comes from this gene: cDNA EST

TITLE
JOURNAL
COMMENT

Anderson, S., Baldwin, J., Barua, N., Beda, F., Boguslavsky, L.,
Boukhalper, B., Brown, A., Burkett, G., Campiolo, A., Castle, A.,
Chopel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
D'Amelio, K., Dewar, K., Douze, S., Donato, M., Doyle, M.,
D'Amelio, K., Dewar, K., Douze, S., Donato, M., Doyle, M.,
Gajana, J., Gajana, S., Gajana, S., Gajana, S., Gajana, S.,
Grand-Pierre, N., Grant, C., Hagos, B., Heald, M., Hearn, D.,
Hewland, J., Iliev, I., Johnson, R., Jones, C., Kain, L., Karatas, A.,
Klein, J., Lander, T., Laroque, K., Lechoczy, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M.,
McKen, P., McGarr, K., McKernan, K., McPherson, R., Melidri, J.,
Merrill, C., Miller, C., Miller, C., Miller, C., Miller, C., Miller, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, A., O'Neil, A.,
Petersen, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., R.,
Riley, R., Rogoy, P., Rothman, D., Roy, A., Santos, R., Scheuer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Sudramanin, A., Talamas, J., Testeyre, S., Theodore, J., Tiroli, A.,
Tuberman, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
Zhang, X.

Submitted (17-Feb-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 10, 2000 this sequence version replaced gi:5984473.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www.wi.mit.edu>
Contact: submissions@genome.wi.mit.edu

Project Information
Center Project name: 6532
Center Project ID: 6532
Sequencing vector: M13: M7815; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161281 bases at least Q40
Consensus quality: 161281 bases at least Q40
Insert size: 162000; approx 66 at least Q20
Insert size: 22151; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
contains 102 contigs. The true order of the pieces
is unknown and the order of the contigs is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved. 1178: contig of 1178 bp in length

1178 1178: contig of 1178 bp in length
1279 1279: contig of 100 bp in length
2326 2425: contig of 1047 bp in length
2326 2425: gap of 100 bp in length
3505 3604: contig of 1079 bp in length
3505 3604: gap of 100 bp in length
3605 4812: contig of 1208 bp in length
4813 4912: gap of 100 bp in length
4813 4912: contig of 424 bp in length
6331 6436: contig of 100 bp in length
6437 7706: contig of 1170 bp in length
7607 7706: gap of 100 bp in length
7707 9220: contig of 1514 bp in length
9221 9320: gap of 100 bp in length
9321 10404: contig of 1084 bp in length
10405 10504: gap of 100 bp in length
10505 11787: contig of 1263 bp in length
11788 11887: gap of 100 bp in length

11888 13396: contig of 1509 bp in length
13397 14496: gap of 100 bp in length
14497 14524: contig of 1028 bp in length
14525 14644: gap of 100 bp in length
14645 16025: contig of 1301 bp in length
16026 16427: contig of 402 bp in length
16428 16527: gap of 100 bp in length
16528 17984: contig of 1457 bp in length
17985 18084: gap of 100 bp in length
18085 19657: contig of 1733 bp in length
19658 19974: gap of 100 bp in length
19975 21462: contig of 100 bp in length
21463 21561: gap of 100 bp in length
21562 23157: contig of 1596 bp in length
23158 23257: gap of 100 bp in length
23258 24526: contig of 1269 bp in length
24527 24626: gap of 100 bp in length
24627 25953: contig of 1329 bp in length
25954 26053: gap of 100 bp in length
26054 27977: contig of 100 bp in length
27978 28076: gap of 100 bp in length
28077 29674: contig of 1728 bp in length
29675 29974: gap of 100 bp in length
29975 31504: contig of 1530 bp in length
31505 31604: gap of 100 bp in length
31605 33072: contig of 1372 bp in length
33073 33423: gap of 100 bp in length
33424 34337: contig of 1161 bp in length
34338 34564: contig of 1317 bp in length
34565 35754: gap of 100 bp in length
35755 37731: contig of 1877 bp in length
37732 39329: contig of 1563 bp in length
39330 41364: contig of 1970 bp in length
41365 41464: gap of 100 bp in length
41465 43336: contig of 1872 bp in length
43337 43436: gap of 100 bp in length
43437 45488: contig of 2052 bp in length
45489 45588: gap of 100 bp in length
45589 47903: contig of 100 bp in length
47904 48022: gap of 100 bp in length
48003 50923: contig of 2921 bp in length
50924 51023: gap of 100 bp in length
51024 53227: contig of 2204 bp in length
53228 53327: gap of 100 bp in length
53329 56857: contig of 3394 bp in length
56858 56871: gap of 100 bp in length
56872 58818: contig of 1867 bp in length
58819 58818: gap of 100 bp in length
58819 61916: contig of 3088 bp in length
61917 62016: gap of 100 bp in length
62017 64893: contig of 2883 bp in length
64894 64999: gap of 100 bp in length
64900 65924: contig of 1477 bp in length
65925 74433: contig of 4857 bp in length
74434 74533: gap of 100 bp in length
74534 78765: contig of 4232 bp in length
78766 78865: gap of 100 bp in length
78866 83783: contig of 4918 bp in length
83784 83883: gap of 100 bp in length
83884 86585: contig of 1671 bp in length
86586 92887: contig of 4233 bp in length
92888 92887: gap of 100 bp in length
92888 97174: contig of 4187 bp in length
97175 97274: gap of 100 bp in length
97275 101602: contig of 4328 bp in length
101603 101702: gap of 100 bp in length
101703 107534: contig of 5632 bp in length

•

•

XX Mouse, growth hormone; GH regulatable gene; liver pathology; hypertrophy;
 XX hepatocellular lesion; hyperplasia; altered expression level; clone 5;
 XX diagnostic marker; gigantism; acromegaly; diabetes; hepatotoxic;
 XX transgenic animal; drug screening; drug discovery; murine; ss.
 XX
 OS Mus sp.
 PN MO200066787-42.
 PN
 PD 09-NOV-2000.
 PD
 XX 05-MAY-2000; 2000MO-US12366.
 XX
 XX 05-MAY-1999; 99US-0132663.
 XX
 XX (UYOH-) UNTV OHIO.
 XX
 XX Kopechek J, Tlone J:
 P1 hpi, 2001-007329,03;
 P2 P-SDS; AA687274, AA68725, AA68726, AA68727, AA68728, AA68729,
 DR AA68730, AA68731, AA68732, AA68733, AA68734, AA68735, AA68736,
 DR AA68737, AA68738, AA68739.
 DR
 XX
 PT Diagnosing abnormal levels of growth hormone activity in liver
 PT comprising assaying growth transcribable activity and protein
 PT expression level of hormone-regulatable liver genes, as diagnostic
 PT markers of liver pathology .
 XX
 XX Claim 1, Page 51-52: 65pp: English.
 XX
 XX The invention relates to a method of diagnosing abnormal levels of
 CC growth hormone (GH) activity in the liver, or predicting a change in the
 CC condition of the liver in response to abnormal GH activity. The method
 CC involves correlating the level of expression of certain specific genes
 CC in the condition of the liver as the result of GH activity. Excessive GH
 CC activity in the liver is thought to be deleterious to health, causing an
 CC increase in liver size as a consequence of both hyperplasia and
 CC hepatocyte hypertrophy, and hepatocellular lesions which progress with
 CC age. Studies in transgenic mice which express high levels of bovine
 CC growth hormone identified a number of genes whose expression in the
 CC liver is altered by high GH levels. The genes which are upregulated are
 CC *IGF1*, *IGF2*, *IGF2R*, *IGFBP3*, *IGFBP5*, *IGFBP6*, *IGFBP7*, *IGFBP8*, *IGFBP9*,
 CC *IGFBP10*, *IGFBP11*, *IGFBP12*, *IGFBP13*, *IGFBP14*, *IGFBP15*, *IGFBP16*,
 CC *IGFBP17*, *IGFBP18*, *IGFBP19*, *IGFBP20*, *IGFBP21*, *IGFBP22*, *IGFBP23*,
 CC *IGFBP24*, *IGFBP25*, *IGFBP26*, *IGFBP27*, *IGFBP28*, *IGFBP29*, *IGFBP30*,
 CC *IGFBP31*, *IGFBP32*, *IGFBP33*, *IGFBP34*, *IGFBP35*, *IGFBP36*, *IGFBP37*,
 CC *IGFBP38*, *IGFBP39*, *IGFBP40*, *IGFBP41*, *IGFBP42*, *IGFBP43*, *IGFBP44*,
 CC *IGFBP45*, *IGFBP46*, *IGFBP47*, *IGFBP48*, *IGFBP49*, *IGFBP50*, *IGFBP51*,
 CC *IGFBP52*, *IGFBP53*, *IGFBP54*, *IGFBP55*, *IGFBP56*, *IGFBP57*, *IGFBP58*,
 CC *IGFBP59*, *IGFBP60*, *IGFBP61*, *IGFBP62*, *IGFBP63*, *IGFBP64*, *IGFBP65*,
 CC *IGFBP66*, *IGFBP67*, *IGFBP68*, *IGFBP69*, *IGFBP70*, *IGFBP71*, *IGFBP72*,
 CC *IGFBP73*, *IGFBP74*, *IGFBP75*, *IGFBP76*, *IGFBP77*, *IGFBP78*, *IGFBP79*,
 CC *IGFBP80*, *IGFBP81*, *IGFBP82*, *IGFBP83*, *IGFBP84*, *IGFBP85*, *IGFBP86*,
 CC *IGFBP87*, *IGFBP88*, *IGFBP89*, *IGFBP90*, *IGFBP91*, *IGFBP92*, *IGFBP93*,
 CC *IGFBP94*, *IGFBP95*, *IGFBP96*, *IGFBP97*, *IGFBP98*, *IGFBP99*, *IGFBP100*,
 CC *IGFBP101*, *IGFBP102*, *IGFBP103*, *IGFBP104*, *IGFBP105*, *IGFBP106*,
 CC *IGFBP107*, *IGFBP108*, *IGFBP109*, *IGFBP110*, *IGFBP111*, *IGFBP112*,
 CC *IGFBP113*, *IGFBP114*, *IGFBP115*, *IGFBP116*, *IGFBP117*, *IGFBP118*,
 CC *IGFBP119*, *IGFBP120*, *IGFBP121*, *IGFBP122*, *IGFBP123*, *IGFBP124*,
 CC *IGFBP125*, *IGFBP126*, *IGFBP127*, *IGFBP128*, *IGFBP129*, *IGFBP130*,
 CC *IGFBP131*, *IGFBP132*, *IGFBP133*, *IGFBP134*, *IGFBP135*, *IGFBP136*,
 CC *IGFBP137*, *IGFBP138*, *IGFBP139*, *IGFBP140*, *IGFBP141*, *IGFBP142*,
 CC *IGFBP143*, *IGFBP144*, *IGFBP145*, *IGFBP146*, *IGFBP147*, *IGFBP148*,
 CC *IGFBP149*, *IGFBP150*, *IGFBP151*, *IGFBP152*, *IGFBP153*, *IGFBP154*,
 CC *IGFBP155*, *IGFBP156*, *IGFBP157*, *IGFBP158*, *IGFBP159*, *IGFBP160*,
 CC *IGFBP161*, *IGFBP162*, *IGFBP163*, *IGFBP164*, *IGFBP165*, *IGFBP166*,
 CC *IGFBP167*, *IGFBP168*, *IGFBP169*, *IGFBP170*, *IGFBP171*, *IGFBP172*,
 CC *IGFBP173*, *IGFBP174*, *IGFBP175*, *IGFBP176*, *IGFBP177*, *IGFBP178*,
 CC *IGFBP179*, *IGFBP180*, *IGFBP181*, *IGFBP182*, *IGFBP183*, *IGFBP184*,
 CC *IGFBP185*, *IGFBP186*, *IGFBP187*, *IGFBP188*, *IGFBP189*, *IGFBP190*,
 CC *IGFBP191*, *IGFBP192*, *IGFBP193*, *IGFBP194*, *IGFBP195*, *IGFBP196*,
 CC *IGFBP197*, *IGFBP198*, *IGFBP199*, *IGFBP200*, *IGFBP201*, *IGFBP202*,
 CC *IGFBP203*, *IGFBP204*, *IGFBP205*, *IGFBP206*, *IGFBP207*, *IGFBP208*,
 CC *IGFBP209*, *IGFBP210*, *IGFBP211*, *IGFBP212*, *IGFBP213*, *IGFBP214*,
 CC *IGFBP215*, *IGFBP216*, *IGFBP217*, *IGFBP218*, *IGFBP219*, *IGFBP220*,
 CC *IGFBP221*, *IGFBP222*, *IGFBP223*, *IGFBP224*, *IGFBP225*, *IGFBP226*,
 CC *IGFBP227*, *IGFBP228*, *IGFBP229*, *IGFBP230*, *IGFBP231*, *IGFBP232*,
 CC *IGFBP233*, *IGFBP234*, *IGFBP235*, *IGFBP236*, *IGFBP237*, *IGFBP238*,
 CC *IGFBP239*, *IGFBP240*, *IGFBP241*, *IGFBP242*, *IGFBP243*, *IGFBP244*,
 CC *IGFBP245*, *IGFBP246*, *IGFBP247*, *IGFBP248*, *IGFBP249*, *IGFBP250*,
 CC *IGFBP251*, *IGFBP252*, *IGFBP253*, *IGFBP254*, *IGFBP255*, *IGFBP256*,
 CC *IGFBP257*, *IGFBP258*, *IGFBP259*, *IGFBP260*, *IGFBP261*, *IGFBP262*,
 CC *IGFBP263*, *IGFBP264*, *IGFBP265*, *IGFBP266*, *IGFBP267*, *IGFBP268*,
 CC *IGFBP269*, *IGFBP270*

[illegible]


```

XX 12-FEB-1999 (first entry)
XX
XX EST clone GP56.
XX
XX Expressed sequence tag; secreted protein; hematopoiesis regulator;
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX chemokinesis; chemokinesis; haemostasis; gene therapy; thrombolytic;
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
XX Homo sapiens.
XX
XX H09845437-A2.
XX
XX 15-OCT-1998.
XX
XX 10-APR-1998: 98W0-US06956.
XX
XX 10-APR-1997: 9705-0837312.
XX
XX (GENB ) GENETICS INST INC.
XX
XX Acostino MJ, Jacobs K, Lawelle ER, McCoy JM, Merberg D:
XX Racle LA, Spaulding V, Treacy M;
XX WPI: 1999-070078/06.
XX
XX New polynucleotide encoding human secreted proteins - derived from
XX human blood cells, including tumour cells, leukocytes, platelets,
XX oratory, pituitary, retina and colon cDNA libraries
XX
XX Claim 1: Page 368: 641pp: English.
XX
XX The present sequence represents an expressed sequence tag (EST), and is
XX a polynucleotide of the invention. The polynucleotide of the invention
XX is a synthetic or natural sequence of nucleotides, which are predicted to
XX encode the EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX hematopoiesis regulating activity, tissue growth activity,
XX haemostatic regulating activity, tumour invasion suppressor activity,
XX and thrombolytic activity, receptor/ligand activity, haemostatic
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX therapy.
XX
XX The EST sequences are also stated to be useful for gene
XX therapy.
XX
XX Sequence 339 BP; 90 A; 74 G; 47 G; 128 T; 0 other:
XX
XX alignment_scores:
XX
XX Quality: 48.00 Length: 14
XX Ratio: 4.364 Gaps: 0
XX Percent Similarity: 78.571 Percent Identity: 50.000
XX
XX alignment_block:
XX US-09-471-2/6-831_COPY_1_16 x AA988355/rev ..
XX
XX Align seq 1/1 to reverse of: AA988355 from: 1 to: 339
XX
XX 2 SerMetLeuValAlaPheLeuLeuLeuTrpGlyValIhrTrp 15
XX ::::::::::::::::::::::::::::::::::::::::::::
XX 74 AGTATATAAATGCTTTTATACACACCTCGGAGGACGGGTGG 33
XX
XX seq name: //SID52/gcgdata/geneseq/geneseq/nh2001.DAT:AA932084
XX
XX seq_documentation_block:
XX ID AA932084 standard: cDNA; 1202 BP.
XX
XX AC AAF32084:
XX
XX 12-APR-2001 (first entry)
XX
XX

```

XX		Human M019 coding sequence.
XX		
XX		Human: M010; M019; M024; cancer; obesity; diabetes; osteoporosis; asthma
XX		chronic obstructive pulmonary disease; hyperlipidemia; arteriosclerosis;
KM		hyperlipoproteinemia; arteriosclerosis; coronary artery disease; sa-
XX		
OS		Homo sapiens.
PN		M0200100544-A1.
XX		
PN		04-JAN-2001.
XX		
PF		30-JUN-2000; 2000MO-US18097.
XX		
PR		30-JUN-1999; 99US-0245293.
XX		
PR		30-JUN-1999; 99US-0245680.
XX		
XX		(MILL-) MILLENNIUM PHARM INC.
XX		
P1		Khododoust M:
XX		
DR		WPI: 2001-112434/12.
XX		
P-		P-PSDB; AAB66893, AAB66894, AAB66895.
XX		
XX		New gene encoding proteins designated M010, M019, M024, useful for the
PT		diagnosis, prophylaxis and treatment of cancer, obesity, diabetes,
XX		osteoporosis, asthma, arteriosclerosis and pulmonary disorders -
XX		
PS		Claim 1; Fig 2; 98pp: English.
XX		
CC		The present invention relates to novel human coding sequences
CC		and proteins designated M010 (AAE32082 and AAB66884), M019 (AAE32084 and
CC		AAB66893) and M024 (AAE32086 and AAB66896). M010, M019 and M024 proteins
CC		bind to and inhibit cellular growth, proliferation, invasion of various cell
CC		lines and/or tissues, induce apoptosis, prevent metastatic dissemination of
CC		cancers, and are useful for the diagnosis, prevention and treatment of a variety of
CC		disorders including cancer, obesity, diabetes, osteoporosis, asthma,
CC		central and peripheral nervous system disorders, pain, stroke, chronic
CC		obstructive pulmonary disease, hypercholesterolemia, hyperlipidemia, .
CC		hyperlipoproteinemia, arteriosclerosis, arterioleclerosis and coronary
CC		artery disease. The present sequence is the coding sequence for human
XX		M019.
XX		
S0		Sequence 1202 BP: 406 A; 226 C; 309 G; 261 T; 0 other;
	alignment_scores:	
	Quality: 48.00	Length: 11
	Ratio: 5.333	Gaps: 0
	Percent Similarity: 81.818	Percent Identity: 72.727
	alignment_block:	
	US-09-471-276-831_COPY_1_16 x AAE32084 ..	
	Align seg 1/1 to: AAE32084 From: 1 to: 1202	
	6 ValPheLeuLeuLeuTrpCysAlaValnThrTrrGly 16	
	TTTTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTT 36	
	364 GTCTTTTAATCTTCAGCGGGCCCTTTATCTGTGGGA 396	
	seq_name: /SINSG/gsgdata/geneseq/geneseqn/NA1998.DAT.AAX13948	
	seq_documentation_block:	
ID	AAX13948 standard; DNA: 1425 BP.	
XX		
AX	AAX13948:	
XX		
DT	31-MAR-1999 (first entry)	
XX		
DE	H. pylori GHPO 1365 geno.	
XX		
NM	GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;	

KM Detection: diagnosis; infection; KP-77-46; KP-85-43; KP-98-22;
 KM KP-98-33; KP-110-32; probe; ds.
 OS Klebsiella pneumoniae.
 XX MO9842843-A1.
 PN 01-OCT-1998.
 PD 23-MAR-1996; 98NO-JP01286.
 PE 25-MAR-1997; 97JP-0071082.
 PR (FUSO) FUSO PHARM IND LTD.
 PA
 XX Abe K, Keshi H, Matsuhisa A, Ueyama H;
 P1 WPI: 1999-045140/04.
 XX
 PR DNA probes for diagnosis of Klebsiella pneumoniae infections - are
 PT obtained by HindIII digestion of genomic K. pneumoniae DNA and
 PT selection of bacterial-binding fragments
 XX
 PS Claim 2: Page 17-18; 32pp; Japanese.
 XX
 CC This sequence is a DNA probe used for the diagnosis of infection with
 CC Klebsiella pneumoniae. The probe is a HindIII restriction enzyme
 CC DNA from this bacterium, digesting it with HindIII to produce
 CC and identifying clones capable of binding to the bacterium then isolating
 CC and sequencing the HindIII fragments involved. The probes have sequences
 CC KP-77-46, KP-85-43, KP-98-22, KP-98-33 and KP-110-32. The probes allow
 CC for simple and highly specific detection of Klebsiella pneumoniae in
 CC biological samples such as blood.
 SO Sequence 1747 BP; 534 A; 367 C; 403 G; 443 T; 0 other;

alignment_scores: Quality: 47.50 Length: 18
 Ratio: 3.167 Gaps: 1
 Percent Similarity: 83.333 Percent Identity: 44.444

alignment_block:
 US-09-471-276-831_COPY_1_16 x AAV71763/rev ..
 Align seg 1/1 to reverse of: AAV71763 from: 1 to: 1747

1 MetSerMetLeuValAlaPheLeuLeu.....TTPGlyValTrp 14
 :::::::::::::::::::::::::::::::::::::::
 1442 GTGACAAATATGATATGATCTGCTGTTGCACATGACCTGGGAAACGAC 1393
 14 TTP 15
 1392 ATCG 1389

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAA74241
 seq_documentation_block:
 ID AAA74241 standard; DNA: 451 BP.
 AC AAA74241;
 XX
 DT 29-NOV-2000 (first entry)
 XX
 DD Lobiolly pine SSR locus R1P185.
 XX
 XX Lobiolly pine: Simple sequence repeat; SSR; microsatellite DNA repeat;
 KW genetic marker; mapping; inheritance study; population genetics study;
 KW plant breeding programme; ss.
 OS Pinus taeda.
 XX
 PM WO200042110-A2.

XX 20-JUL-2000.
 XX 06-JAN-2000; 2000MO-US00325.
 XX
 XX 15-JAN-1999; 99US-0232884.
 PE 19-JAN-1999; 99US-0232785.
 XX
 BA (INTO) INT PAPER CO.
 DA (BCHR) ECHT C S.
 DA (NELS) NELSON C D.
 DA (USDA) US SEC OF AGRIC.
 XX
 PI ECHT CS, Nelson CD;
 XX
 DR WPI: 2000-482836/42.
 XX
 PR Polynucleotide having simple sequence repeat useful as markers in
 PT plant breeding studies for genetic characterization e.g. genetic mapping study, or
 PT inheritance study of a commercially important trait in a plant breeding
 PT program ..
 XX
 PS Claim 1: Pages 32-33; 57pp; English.
 XX
 CC The present sequence is a lobiolly pine simple sequence repeat (SSR)
 CC locus. SSRs are also known as microsatellite DNA repeats. The present
 CC sequence is useful as a genetic marker for genetic mapping, population
 CC genetic studies and inheritance studies in various plant breeding
 CC programmes.
 SO Sequence 451 BP; 209 A; 57 C; 63 G; 122 T; 0 other;

alignment_scores: Quality: 47.00 Length: 13
 Ratio: 3.615 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 53.846

alignment_block:
 US-09-471-276-831_COPY_1_16 x AAA74241/rev ..
 Align seg 1/1 to reverse of: AAA74241 from: 1 to: 451

3 MetLeuValAlaPheLeuLeuTrpGlyValTrp 15
 :::::::::::::::::::::::::::::::::::::::
 165 TTATPATTGTTCTGTTGTTATGAGGACATPACATAC 127

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAC87258
 seq_documentation_block:
 ID AAC87258 standard; CDNA: 1453 BP.
 AC AAC87258;
 XX
 DT 09-MAR-2001 (first entry)
 XX
 DD Mouse liver growth hormone-induced cDNA clone 45, SEQ ID NO:9.
 XX
 XX Mouse: growth hormone; GH regulatable gene; liver pathology; hypertrophy;
 KW hepatocellular lesion; hyperplasia; altered expression level; clone 45;
 KW diagnostic marker; gigantism; acromegaly; diabetes; hepatotropic;
 KW transgenic animal; drug screening; drug discovery; mouse; murine; ss.
 OS Mus sp.
 XX
 PM WO2000066787-A2.
 XX
 PD 09-NOV-2000.
 PF 05-MAY-2000; 2000MO-US12366.
 XX
 XX 05-MAY-1999; 99US-0132663.
 XX


```

XX cervical cancer; ss.
OS Homo sapiens.
PN M0200157278-A2.
PD 09-AUG-2001.
PX 30-JAN-2001; 2001WO-US00670.
PF 04-FEB-2000; 2000US-0189312.
PI 06-MAY-2000; 2000US-0189312.
PR 30-JUN-2000; 2000US-0608448.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234697.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024253.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Hank DR:
DX MPI: 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX claim 25; SEQ ID NO 4735; 487pp; English.
XX The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC cancer.
CC Note: The sequence data for this patent did not come from part of the printed
CC specification. It was obtained in electronic format directly from MIPO
CC at ftp:wipo.int/pub/published_patl_sequences.
XX
XX Sequence 1566 BP; 406 A; 402 C; 610 G; 548 T; 0 other:
SQ
alignment_scores:
Quality: 46.00 Length: 22
Ratio: 3.067 Cases: 1
Percent Similarity: 68.182 Percent Identity: 36.364
alignment_block:
US-09-471-276-B31_COPY_1-16 x AA114802 ..
Align seq 1/1 to: AA114802 from: 1 to: 1566
1 MetSerMetLeuValVal... Phalaenautaurf 11
|||||.....|||.....|||.....|||
1282 ATGACACTTCATGGAGCCTGCAGGTGAACAATCGTCACTCATGATGTCG 1331
11 polyA tailTTTTCpCpY 16
1332 GAGCGTGACCTGAGCGGT 1347
seq_name: /SID52/gcgcdbdata/geneseq/geneseqn/NA2001.DAT:AA136160
seq_documentation_block:
ID AA136160 standard; DNA: 1566 BP.
XX AA136160:
DX 17-OCT-2001 (first entry)
XX Probe #4846 used to measure gene expression in human placenta sample.
XX Probe: microarray: human; placenta; antenatal diagnosis:
XX

```

```

XX genetic disorder; ss.
OS Homo sapiens.
XX
XX M0200157272-A2.
XX
PD 09-AMG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 04-FEB-2000; 2000US-0684406.
XX
PR 30-MAY-2000; 2000US-0684406.
XX
PR 03-AMG-2000; 2000US-0632365.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000G6-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SQ, Hanzel DK, Chen W, Rank DR.
XX
XX WPI; 2001-488897/53.
XX
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta.
XX
XX
XX Claim 25; SEQ ID NO 4846; 654bp; English.
XX
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are displaying for
XX CC producing a microarray for predicting, measuring and analyzing gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 1566 BP; 406 A; 402 C; 610 G; 548 T; 0 other:
XX
XX
XX alignment_scores:
XX      Quality: 46.00      Length: 22
XX      Ratio: 3.067      Gaps: 1
XX      Percent Similarity: 68.182      Percent Identity: 36.364
XX
XX alignment_block:
XX US-09-471-276-831_COPY_1_16 x AA1045160 ..
XX
XX Align seq 1/1 to: AA136160 from: 1 to: 1966
XX
XX      1 MetSerMetLeuValVal.....PheLeuLeuLeuTrp 11
XX      |||||.....|||||.....|.....|.....|.....|
XX      1282 ATGAGCTTAAAGTACCTGACGCTGAACACTGACTGTCATGTCATGTCG 1331
XX      || polyaattatgtrficy 16
XX      |.....|.....|.....|
XX      1332 GAGCTGACCTGCGGCGT 1347
XX
XX seq_name: /S1052/9c9data/geneseq/geneseq/NA2001.DAT:AA104592
XX
XX seq_documentation_block:
XX      AA104592 standard; DNM; 1566 BP.
XX
XX AA104592:
XX
XX 09-OCT-2001 (first entry)
XX
XX Probe #4583 used to measure gene expression in human breast sample.
XX
XX Probe: human; breast disease: breast cancer; development disorder; sci
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour
XX
XX Homo sapiens.
XX
XX M0200157270-A2.

```

```

XX 09-AUG-2001.
XX 29-JAN-2001: 2001WO-US000661.
XX 04-FEB-2000: 2000US-0189312.
XX 26-MAY-2000: 2000US-0287454.
XX 30-JUN-2000: 2000US-0608408.
XX 03-AUG-2000: 2000US-0632366.
XX 21-SEP-2000: 2000US-0234687.
XX 27-SEP-2000: 2000US-0236359.
XX 04-OCT-2000: 2000CB-0024283.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX WPI: 2001-476286/51.
XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX Claim 25: SEQ ID NO 4583; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes.
XX The present sequence is one such probe. The probes are useful for
XX measuring human gene expression in a human breast sample, where the probe
XX hybridises at high stringency to a nucleic acid expressed in the human
XX breast. The probes are useful for predicting, diagnosing, grading,
XX screening, and prognosticating human breast cancer, and for the treatment
XX particularly those diseases with polygenic aetiology. The diseases
XX include: breast cancer, disorders of development, inflammatory diseases
XX of the breast, fibrocystic changes, proliferative breast disease and
XX non-carcinoma tumours.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at http://wipo.int/pub/published\_pcc\_sequences.
XX Sequence 1966 BP: 406 A; 402 C; 610 G; 548 T; 0 other:
XX
XX Alignment_scores:
XX Quality: 46.00 Length: 22
XX Ratio: 4.162 Gaps: 0
XX Percent Similarity: 68.182 Percent Identity: 36.364
XX
XX alignment_block:
XX US-09-471-276-831_COPY_1.16 x AA104592 ..
XX
XX Align seg 1/1 to: AA104592 from: 1 to: 1966
XX
XX 1 MetSeMetLeuValVal.....PheGluLeuTrp 11
XX |||||.....|
XX 1332 GAGGCTGACCTGGGGT 1347
XX
XX seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1599.DAT:AA17774
XX
XX seq_documentation_block:
XX ID AA17774 standard; cDNA: 2394 BP.
XX
XX AA17774:
XX AC
XX 12-MAY-1999 (first entry)
XX
XX YAK-1: serine-threonine protein kinase-RTKAR33 encoding cDNA.
XX inflammatory disorder; osteoporosis; Adult Respiratory Disease Syndrome;
XX arthritis; psoriasis; dermatitis; asthma; allergy; infection; HIV-1;

```

```

XX HIV-2; cachexia; immunodeficient disorder; septic shock; pain; injury;
XX cancer; anorexia; bulimia; Parkinson's disease; cardiovascular disease;
XX neurological disorder; Huntington's disease; gene therapy; gene mapping;
XX Gilles de la Tourette's syndrome; ss.
XX
XX Homo sapiens.
XX
XX E8894863.AL.
XX
XX 03-FEB-1999.
XX
XX 21-JUL-1998: 98EP-0305794.
XX
XX 20-FEB-1999: 98US-0027064.
XX 28-JUL-1997: 97US-0053924.
XX
XX (SMK ) SMITHKLINE BEECHAM CORP.
XX
XX Bergsma DJ, Shabon U:
XX WPI: 1999-108353/10.
XX P-PSDB: AAN94997.
XX
XX New serine-threonine kinase (HTKAR33) polypeptides and
XX polynucleotides - useful as diagnostic reagents and for prevention
XX and treatment of bone loss, neurological and inflammatory disorders
XX and cancer, HIV infections and angina pectoris
XX
XX Claim 2: Page 13-15; 30pp; English.
XX
XX This cDNA encodes a YAK-1 related serine-threonine protein kinase
XX polypeptide (HTKAR33). Host cells transformed with an expression vector
XX comprising the HTKAR33 nucleic acid are used for the recombinant
XX production of the protein. HTKAR33 polynucleotides and polypeptides are
XX useful for are useful for diagnosing susceptibility to diseases and for
XX screening for antagonists, agonists. These can be used in treatment to
XX against bone loss, osteoporosis, HIV-1, HIV-2, HIV-related diseases
XX and disorders including osteoporosis, Adult Respiratory Disease Syndrome
XX (ARDS), Rheumatoid arthritis, osteoarthritis, inflammatory Bowel Disease
XX (IBD), psoriasis, dermatitis, asthma, allergies; bacterial, fungal,
XX protozoan and viral infections, especially those caused by HIV-1 or
XX HIV-2; HIV-associated cachexia and other immunodeficient disorders;
XX septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's
XX disease; Huntington's disease; cardiovascular disease; gene therapy;
XX acute heart failure; myocardial infarction; hypotension; hypertension;
XX urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy;
XX and psychotic and neurological disorders, including schizophrenia, manic
XX depression, anxiety; delirium; dementia; severe mental retardation and
XX dyskinesias, such as Huntington's disease or Gilles de la Tourette's
XX syndrome. HTKAR33 polypeptides are also useful for mapping genes to
XX chromosomes, allowing gene inheritance to be studied through linkage
XX analysis.
XX
XX Sequence 2394 BP: 478 A; 801 C; 676 G; 439 T; 0 other:
XX
XX Alignment_scores:
XX Quality: 46.00 Length: 14
XX Ratio: 4.162 Gaps: 0
XX Percent Similarity: 76.571 Percent Identity: 57.143
XX
XX alignment_block:
XX US-09-471-276-831_COPY_1.16 x AA17774 ..
XX
XX Align seg 1/1 to: AA17774 from: 1 to: 2394
XX
XX 2 SerMetLeuValValPheGluLeuTrpGlyValThrTrp 15
XX |||||.....|
XX 1901 TCCGCTCTCTGATACCTGCACCTGCGGGGCTGCACCTGG 1942
XX
XX seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA2000.DAT:AA16647

```

[seq_documentation_block:](#)

ID AA016647 standard; cDNA, 3191 BP.

XX

DT 16-JUN-2000 (first entry)

XX

DE Human secreted protein clone ck213_12 nucleotide sequence S80 id NO:59.

XX

KW Human; secreted protein clone; immunestimulant; immunesuppressant; vtrinicide;

KW antibacterial; antitumoral; cytostatic; antiinflammatory; dermatological;

KW antiviral; immune deficiency; severe combined immunodeficiency; SCID;

KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;

KW connective tissue disease; multiple sclerosis; erythematosis;

KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;

KW guillain-barre syndrome; autoimmune thyroiditis; myasthenia gravis;

KW insulin dependent diabetes mellitus; graft-versus-host-disease;

KW autoimmune inflammatory eye disease; allergy; ss.

CX Homo sapiens.

PX MO2000095552-A1.

NN

PD 24-FEB-2000.

XX

L3-AUG-1999; 99MO-USI8298.

XX

XZ 17-AUG-1998; 98US-00966522.

PR 17-AUG-1998; 98US-00968185.

PR 04-SEP-1998; 98US-00932289.

PR 28-OCT-1998; 98US-01053619.

PR 08-JAN-1999; 99US-01152324.

PR 12-FEB-1999; 99US-01199311.

PR 16-MAR-1999; 99US-01230070.

PR 10-Apr-1999; 99US-00966522.

XX

(CEMY) GENETICS INST INC.

PA Jacobs K., McCoy JM., Lavallee ER., Collins-Racie LA., Evans C.;
P1 McRory D., Treacy M., Apostino MJ., Steininger RJ., Spaulding V.;
PT Meng CC., Clark HF., Fechtel R;
DR WPI: AAA949597./1B.
PX P-FSDb: 2000-205979./1B.

CC New polynucleotides encoding proteins, which may have e.g.
CC nutritional, chemone, immune stimulating or suppressing,
CC hematopoiesis regulating, tissue growth, activator/inhibitor
CC immunoinflammatory or tumor inhibition activity -
PS Claim 68: Page 530-531; 641pp: English.

AA016618 to AA016697 encode the human secreted proteins given in
AAAY94968 to AAAY94980. Isolated from human adult brain, adult thyroid,
adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
adult placenta, adult testis, whole embryo, adult cartilage, kidney,
lung, adult bladder, cDNA libraries. The polynucleotides and proteins are
predicted to have biological activities which would make them suitable
for treating, preventing or ameliorating medical conditions in humans
and animals. The polynucleotides can be used as markers for tissues in
which the protein is preferentially expressed, as molecular weights
markers on Southern gels, and as chromosome makers or tags to identify
chromosomes or to map gene positions. The proteins can be used in the
treatment of diseases such as severe combined immunodeficiency (SCID), as well as viral infections. These infections include human immunodeficiency virus (HIV),
hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
candidiasis. The proteins can be used to treat autoimmune disorders such
as connective tissue disenesr, multiple scleriosis, systemic lups
erythematosis, rheumatoid arthritis, autoimmune pulmonary inflammation,

	CC	Gullibila-barre syndrome, autoimmune thyroiditis, insulin-dependent diabetes mellitus, myasthenia gravis, graft-versus-host-disease and autoimmune inflammatory eye disease. The proteins can also be used to treat allergic conditions, such as asthma.
	CC	AAL0698 to AAL0774 represent probes for the humen secreted Proteasins from the present invention..
	CX	Sequence 3191 Pst; 99B A: 716 C: 835 G: 642 T: 0 other;
XX	SQ	
XX	alignment_scores:	
XX	Quality:	46.00 Length: 13
XX	MATCH:	76.923 Caps: 0
XX	Percent Similarity:	Percentage Identity: 61.538
XX	Alignment Block:	
XX	US-09-471-276-B31.COPY_1.16 x AAL16647 ..	
XX	Align seg 1/1 to: AAL16647 from: 1 to: 3191	
XX	4 LevalValPheLeuleutrcPolYnAlThrTrGly 16	
XX	2840 TTACCTCCTTTTCATCGATTGTGCCTGGTGGSGTGGSAG 2878	
seq_name: /SIDSE/gcgbda/g/enaseq/geneseqn/AAL2001.DMT:AALH54911		
ID: AALH54911 standard: DNA: 3427 BP.		
Pt AA HAH54911:		
Dt XX 03-sep-2001 (first entry)		
XN DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:4275.		
KM Staphylococcus epidermidis SKI strain: Infection: diagnosis:		
OS Vaccination: endocarditis : ds.		
SS Staphylococcus epidermidis-		
PN MO200134809-A2.		
PD 17-MAY-2001.		
PR 09-nov-2000: 2000MC-US50782.		
PS 09-nov-1999: 99NS-0164258.		
PA GLAXO) GLAXO GROUP LTD.		
PI Kimmertly MJ:		
NF1: 2001-316495/33.		
NI: nucleic acids encoding polypeptides from Staphylococcus epidermidis,		
NU: useful for vaccinating against infections, e.g. endocarditis -		
Claim B: Page 2004-2005: 2188pp: English.		
AAHS3204 to AHAS3970 represent nucleic acids (I) encoding polypeptides		
(II), or antibodies specific to said polypeptides (III). Said polypeptides		
(II) and (III) may have antibacterial activity and therefore can be used		
in vaccination. The nucleic acid (I) may be used to produce the		
S. epidermidis polypeptides (II) via the production of vectors		
containing them which are used to produce hosts coles which express		
the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be		
used to vaccineate subjects and to raise antibodies against the bacteriae		
treatment of S. epidermidis infection.		
AAHS3970 to AAHS5090 represent specifically claimed S. epidermidis genomic DNA		
polynucleotide sequences from the present invention. AHA5091 to		
AHH5098 represent oligonucloetide sequences and primers which are used		
in the exemplification of the present invention.		

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3427 BP: 1167 A; 642 C; 478 G; 1140 T; 0 Other:

alignment_scores:

Quality:	46.00	Length:	15
Ratio:	3.538	Gaps:	0
Percent Similarity:	86.667	Percent Identity:	40.000

alignment_block:

US-09-471-276-831_COPY_1_16 x AAH54911/rev ..

Align seq 1/1 to reverse of: AAH54911 from: 1 to: 3427

1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrp 15

1442 GTGACATCCCTGCTGATGATGATGATGATGATGATGATGATG 1398

seq_name: /SID52/gcgcdata/geneseq/geneseq/NM2001.DAT:AAH54481

seq_documentation_block:

ID AAH5481 standard: DNA: 3450 BP.

AAH5481:

03-SEP-2001 (first entry)

S. epidermidis genomic polynucleotide sequence SEQ ID NO:3845.

Staphylococcus epidermidis SRI strain: Infection; diagnosis:

vaccination; endocarditis; ds.

Staphylococcus epidermidis.

MO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000MO-US30782.

09-NOV-1999; 99US-0164258.

(GLAXO) GLAXO GROUP LTD.

Kimmerly MJ;

WPI: 2001-316495/33.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis -

Claim 8: Page 1487-1488; 2188pp: English.

AAH53970 represent nucleic acids (1) encoding polypeptides (11), given in AAC81454 to AAC81320, from Staphylococcus epidermidis (11) which can have antibacterial activity and therefore can be used in vaccination therapy against the production of vectors.

S. epidermidis polypeptides (11) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (11) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of infections.

AAH55090 represent specifically claimed S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3450 BP: 1206 A; 679 C; 486 G; 1079 T; 0 Other:

alignment_scores:

Quality:	46.00	Length:	15
Ratio:	3.538	Gaps:	0
Percent Similarity:	86.667	Percent Identity:	40.000

alignment_block:

US-09-471-276-831_COPY_1_16 x AAH54481/rev ..

Align seq 1/1 to reverse of: AAH54481 from: 1 to: 3450

1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrp 15

2470 GTGACATCCCTGCTGATGATGATGATGATGATGATGATGATG 2426

seq_name: /SID52/gcgcdata/geneseq/geneseq/NM2001.DAT:AAH54639

seq_documentation_block:

ID AAH54639 standard: DNA: 3599 BP.

AAH54639:

03-SEP-2001 (first entry)

S. epidermidis genomic polynucleotide sequence SEQ ID NO:4003.

Staphylococcus epidermidis SRI strain: Infection; diagnosis:

vaccination; endocarditis; ds.

Staphylococcus epidermidis.

MO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000MO-US30782.

09-NOV-1999; 99US-0164258.

(GLAXO) GLAXO GROUP LTD.

Kimmerly MJ;

WPI: 2001-316495/33.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis -

Claim 8: Page 1677-1678; 2188pp: English.

AAH53970 represent nucleic acids (1) encoding polypeptides (11), given in AAC81454 to AAC81320, from Staphylococcus epidermidis (11) which can have antibacterial activity and therefore can be used in vaccination therapy against the production of vectors.

S. epidermidis polypeptides (11) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (11) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of infections.

AAH55090 represent specifically claimed S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

CC N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequences are given in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present for SEQ ID NO:4453 to 4464.

XX Sequence 3599 BP; 1144 A; 567 C; 692 G; 1196 T; 0 other:

alignment_scores:

Quality:	46.00	Length:	15
Ratio:	3.528	Gaps:	0
Percent Similarity:	86.667	Percent Identity:	40.000

alignment_block:

us-09-471-276-831_COPY_1.16 x AAH54639 ..

Align seg 1/1 to: AAH54639 from: 1 to: 3599

1 Mot:selevalvalphleuleuleutrfglyvalthrptp 15

1884 GTGACATTCCTCATTCATTCCTTCATTCGCGGCTTTATGC 1928

seq_name: /SID52/gcgcdata/geneseq/geneseqn/NA2000.DNT:AACT6440

seq_documentation_block:

ID AACT6440 standard; cDNA; 4522 BP.

AC AACT6440:

DT 08-FEB-2001 (first entry)

XX Human OREX ORF1995 polynucleotide sequence SEQ ID NO:3989.

XX Human; open reading frame; OREX; detection: cytoethic; hepatotropic; XX vulnerability: antipneumatic; antiparkinson; neurotropic; neuroprotective; XX anticonvulsant; osteopathic; antilethritic; immunosuppressant; cridant; XX immunostimulant; thrombolytic; coagulant; vasotropic; antidabetic; XX hypotensive; dermatological; immunosuppressive; antinflammatory; XX antiviral; antibacterial; antitumor; antineumatic; antilethroid; XX antineumatic; gene therapy; cancer; proliferative disorder; hypertension; XX cardiovascular disease; diabetes mellitus; erythematous; AIDS; XX cholesteral ester storage; systemic lupus erythematosus; infection; XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma; XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; XX bone damage; cartilage damage; coagulation; XX thrombosis; contraceptive; ss.

OS Homo sapiens.

PN W0200058473-A2.

PD 05-OCT-2000.

XX 31-MAR-2000; 2000NO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

PA (CURA-) CURANG CORP.

PA Shinkens RA, Leach M;

DR WPI: 2000-602362/57.

PR P-PSDB: ABA42231.

XX Novel nucleic acids and peptides derived from open reading frame X, XX useful for treating e.g. cancers, proliferative disorders, XX neurodegenerative disorders and cardiovascular disease -

PS Claim 5: Page 3161-3164; 5507pp; English.

XX AACT7446 to AACT7606 encode the proteins given in ABA40237 to ABA4397, CC which represent the human OREX open reading frames 1 to 3161. The OREX CC sequences have activities such as: cytosolic; hepatotropic; vulnerability; CC osteopathic; anticonvulsant; antilethritic; immunosuppressant; CC immunostimulant; cridant; thrombolytic; coagulant; vasotropic; CC antidiabetic; hypotensive; dermatological; immunosuppressive; CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic; CC antilethroid; and antineumatic. The sequences can be used for determining CC the presence of or predisposition to, or preventing or treating CC the presence of or predisposition to, an OREX-associated disorder. The CC polynucleotides can be used for the treatment of the disorder. CC vectors. The proteins and nucleic acids may be used to treat cancers, CC proliferative disorders, neurodegenerative disorders, osteoarthritis, CC graft vs host disease, cardiovascular disease, diabetes mellitus, CC hypertension, hypothyroidism, cholesteral ester storage, systemic lupus CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, CC bacterial or fungal infection, malaria, autoimmune disorders, asthma, CC myalgia, aplastic anaemia, thrombosis, bone and cartilage damage, CC osteoarthritis, rheumatoid arthritis, and as a contraceptive. CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 4522 BP; 1294 A; 1006 C; 1144 G; 1076 T; 2 other:

alignment_scores:

Quality:	46.00	Length:	13
Ratio:	4.600	Gaps:	0
Percent Similarity:	76.923	Percent Identity:	61.538

alignment_block:

us-09-471-276-831_COPY_1.16 x AACT6440 ..

Align seg 1/1 to: AACT6440 from: 1 to: 4522

4 leuvalvalphleuleuleutrfglyvalthrptp 16

3313 TTAGCTGTTTTCCTCATTCCTTCCTTCATTCGCGGCTTTATGC 3351

seq_name: /SID52/gcgcdata/geneseq/geneseqn/NA2001.DNT:AAH9053

seq_documentation_block:

ID AAH9053 standard; cDNA; 5345 BP.

AC AAH9053:

DT 12-OCT-2001 (first entry)

XX Human EST-derived coding sequence SEQ ID NO: 910.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;

XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;

XX diagnostics; forensic test; gene mapping; genetic disorder;

XX biodiversity; gene therapy; nutrition; ss.

OS Homo sapiens.

PN W0200154477-A2.

PD 02-AUG-2001.

XX 25-JAN-2001; 2001NO-US02687.

XX 25-JAN-2000; 2000US-0491404.

PR 17-JUN-2000; 00US-0617746.

PR 03-AUG-2000; 00US-0631431.

PR 15-SEP-2000; 2000US-063870.

PA (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

DR P-PSDB; AAW79529.
 PR Novel mouse and human circadian rhythm gene, clock - useful for
 PT treating e.g., jet-lag, sleep-wake disorders, abnormal cell division,
 PP etc.
 PS Claim 4; Fig 8.1-8.8; 154pp; English.
 XX
 CC This is the complete nucleotide sequence of the murine Clock gene
 CC based upon genomic exon sequences; splice donor and acceptor sites
 CC are provided in AAW6103-48. It encodes a 885-amino acid polypeptide
 CC (p100). The protein has a transcription factor, the Clock gene
 CC regulates at least 2 fundamental circadian clock genes, the Clock
 CC system; the intrinsic circadian period and the persistence of
 CC circadian rhythmicity. Identification and isolation of the Clock
 CC gene involved: genetic mapping of the mouse genome (mid portion of
 CC chromosome 5); Physical mapping of the Clock region; transfection
 CC unit analysis in the Clock region by (a) direct screening of SCN
 CC cDNA libraries with BAC clones as probes, (b) hybridisation
 CC and (c) shuttle sequencing random M13 libraries. BAC clones, BAC
 CC clones, transgenic mouse expression of an isolated BAC clone and
 CC phenotypic rescue of Clock; and mRNA expression of candidate genes
 CC by Northern analysis in Clock mutant versus wild-type mice. The
 CC human Clock gene (see AAW61450) has also been identified. The
 CC invention provides isolated and purified Clock polypeptides. The
 CC polypeptides (including antisense), vectors and host cells. The
 CC polypeptides are useful for the treatment of disorders of the
 CC altered or disrupted circadian rhythms e.g., jet-lag, seasonal
 CC affective disorder, sleep-wake cycle disorders such as mood state,
 CC stress, neurological disorders, to regulate diet and food intake
 CC especially for diabetes, to treat cardiovascular, respiratory,
 CC liver or endocrine disorders, and for diagnosis and treatment of
 CC abnormal cell division such as cancer.
 XX
 SQ Sequence 7498 BP; 2197 A; 1534 C; 1616 G; 2151 T; 0 other;
 XX
 Alignment_scores:
 Quality: 46.00 Length: 9
 Ratio: 5.111 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 88.889
 Alignment_block:
 US-09-471-276-831_COPY_1_16 x AAW61401 ...
 Align seg 1/1 to: AAW61401 from: 1 to: 7498
 8 LeukoleukTTPGlyValThrTTPGly 16
 5598 CTCCTCTCAAGGCGTCAATTCGGCT 5624
 seq_name: /SID52/gcdata/geneseq/geneseq/NA2001.DAT:AA121916
 seq_documentation_block:
 ID AA121916 standard; DNA; 128 BP.
 XX AA121916:
 XX
 DT 12-OCT-2001 (first entry)
 DE Probe #11849 for gene expression analysis in human cervical cell sample.
 KW Probe: human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-050670.

XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-060408.
 PR 23-SEP-2000; 2000US-0234289.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PA Penn SC, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 DR
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 25; SEQ ID No 11849; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SNPs). The present sequence is one such probe. The SNPs are derived
 CC from human HeLa cells. The SNPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC The present invention also provides a part of the part of the printed
 CC specification but was obtained from electronic format directly from WPI
 CC at ftp.wpi-inc/pub/published_pcl_sequences.
 XX
 SQ Sequence 128 BP; 18 A; 40 C; 37 G; 33 T; 0 other;
 XX
 Alignment_scores:
 Quality: 45.00 Length: 11
 Ratio: 5.000 Gaps: 0
 Percent Similarity: 81.818 Percent Identity: 63.636
 Alignment_block:
 US-09-471-276-831_COPY_1_16 x AA121916 ...
 Align seg 1/1 to: AA121916 from: 1 to: 128
 5 ValValPheLeuLeuLeuLeuTTPGlyValThrTTP 15
 74 GTCTCTTTGTCCTCTGCGGCTGCGGCTGCGG 106
 seq_name: /SID52/gcdata/geneseq/geneseq/NA2001.DAT:AA147201
 seq_documentation_block:
 ID AA147201 standard; DNA; 128 BP.
 XX AA147201:
 XX
 DT 17-OCT-2001 (first entry)
 DE Probe #15867 used to measure gene expression in human placenta sample.
 KW Probe: microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-050663.
 XX
 PD 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-060408.
 PR 23-SEP-2000; 2000US-0234289.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

PR	03-AUG-2000;	2000US-0632365.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR:	
XX		
DR	WPI: 2001-48897/53.	
XX		
PT	Human genome-derived single exon nucleic acid probes useful for	
XX	analyzing gene expression in human placenta	
PS	Claim 25; SEQ ID NO 2772; 654bp; English.	
CC	The present invention relates to single exon nucleic acid probes (SENP).	
CC	The present sequence is one such probe. The probes are useful for	
CC	producing a microarray for predicting, measuring and displaying gene	
CC	expression in samples derived from human placenta. The probes are useful	
CC	for prenatal diagnosis of human genetic disorders.	
SQ	Sequence 472 BP: 102 A: 121 C: 145 G: 104 T: 0 other:	
	alignment_scores:	
	Quality: 45.00 Length: 11	
	Ratio: 5.000 Gaps: 0	
	Percent Similarity: 61.818 Percent Identity: 63.636	
	alignment_block:	
	US-09-471-276-831_COPY_1.16 x AA134086 ..	
	Align seg 1/1 to: AA134086 from: 1 to: 472	
	5 ValValPheGluLeuTTCtClValHisTrp 15	
	: : :	
	380 GTCTCATTGTCCGCGCTGTGGGGCTGTGGGTGG 412	
	seq_name : /SIDS2/gcgbdata/geneseq/genaseqn/NA2001.DAT:AA102644	
	seq_documentation_block:	
ID	AA102644 standard; DNA: 472 BP.	
XX		
XX	AA102644:	
XX		
DT	09-OCT-2001 (first entry)	
DE	Probe #2635 used to measure gene expression in human breast sample.	
KM	Probe: human; breast disease: breast cancer; development disorder; as;	
KM	Inflammatory disease: proliferative breast disease; non carcinoma tumour	
OS	Homo sapiens.	
XX		
FN	WO200157270-A2.	
XX		
PD	09-AUG-2001.	
XX		
XX	29-JAN-2001; 2001WO-US00661.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632365.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000GB-0024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR:	
XX		
PT		


```

XX (EVOT-) EVOTEC BIOSYSTEMS AG.
XX Goetz F., Pohnler J., Strauss A., Thumm G.
XX MPI: 1999-255102/21.
XX Identifying agents that affect adhesion of Gram positive bacteria
XX
XX Disclosure: Page 38-39: 51pp; German.
XX This invention describes a novel method for identifying an active agent
XX that affects covalent bonding of a protein to the surface of
XX Gram-positive bacteria. The agents identified inhibit
XX polypeptide-induced adhesion of bacteria to extracellular matrix in the
XX host. The active agent affects the LPXTG motif-dependent anchorage of
XX pathogenically factors to the bacterial cell wall and are potentially
XX useful as therapeutic agents, inhibiting binding of bacteria to
XX extracellular matrix in the host.
XX The agents described in the specification but is represented
XX in the sequence ID listing.
XX
XX Sequence 748 BP: 216 A: 113 C: 170 G: 238 T: 11 other:
XX
XX Alignment_scores:
XX Quality: 45.00 Length: 8
XX Ratio: 6.428 Gaps: 0
XX Percent Similarity: 87.500 Percent Identity: 87.500
XX
XX alignment_block:
XX us-09-471-276-831_copy_1_16 x AA024250 ..
XX
XX Align seg 1/1 to: AA024250 from: 1 to: 748
XX
XX      9 LeuLeuTrpGlyValThrTrpGly 16
XX      |||||
XX      585 CTTTGTGCGGCCGCCACCTGGGCT 608
XX
XX seq_name: /SID52/gcdata/geneseq/geneseq/NA1995.DAT:AA087927
XX
XX seq_documentation_block:
XX ID AA087927 standard: DMT, 794 BP.
XX
XX AC AA087927:
XX
XX DT 22-NOV-1995 (first entry)
XX
XX DX Anthr specific gene promoter element.
XX
XX RX Anthr specific gene: Brassica napus; antisense RNA; Cruciferae;
XX pollen formation; male sterile plant; ss.
XX
XX CS Brassica napus.
XX
XX XX JP07059573-A.
XX
XX PD 07-MAR-1995.
XX
XX PE 20-AUG-1993: 93JP-0206459.
XX
XX PR 20-AUG-1993: 93JP-0206459.
XX
XX RA (MITS) MITSUBISHI CORP.
XX
XX RA (MITO) MITSUBISHI KASEI CORP.
XX
XX DR MPI: 1995-135897/18.
XX
XX PT A gene expressed specifically in anther(s) - used for the
XX preparation of male sterile Cruciferae plants
XX
XX Claim 5: Page 5-6: 6pp; Japanese.
XX

```

```

CC This sequence was used for the expression of an anther specific gene
CC derived from Brassica napus. This sequence caused expression of
CC the anther specific gene as antisense RNA in a Cruciferae plant causing
CC pollen formation to be modified. This is esp. useful in the production
CC of male sterile plants.
XX
XX Sequence 794 BP: 258 A: 157 C: 137 G: 228 T: 14 other:
XX
XX Alignment_scores:
XX Quality: 45.00 Length: 15
XX Ratio: 3.750 Gaps: 0
XX Percent Similarity: 80.000 Percent Identity: 46.667
XX
XX alignment_block:
XX us-09-471-276-831_copy_1_16 x AA087927/rev ..
XX
XX Align seg 1/1 to reverse of: AA087927 from: 1 to: 794
XX
XX      2 SerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGly 16
XX      .....
XX      763 TCTTCTTGTGTGATTTTTCGTGCTGTATGACCCTTTATGAGGCG 719

```


A1BQ_HUMAN STANDARD: 474 AA.
 ID P04217
 DT 20-MAR-1987 (rel. 04, Created)
 DT 20-MAR-1987 (rel. 04, Last sequence update)
 DE 20-AUG-2001 (rel. 40, Last annotation update)
 DE ALPHA-1B-GLYCOPROTEIN.
 OS A1BQ.
 OS Homo sapiens (human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eulacerta; Primates; Catarrhini; Hominoidea; Homo.
 CC [1]_TaxID:9606.
 RN SEQUENCE, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RA MEDLINE:86205955; PubMed:3458201;
 RA Ishioka N., Takahashi N., Putnam F.W.;
 RA "Amino acid sequence of human plasma alpha 1b-glycoprotein: homology
 RA to the immunoglobulin supergene family."
 RA Proc. Natl. Acad. Sci. U.S.A. 83:285-287(1986).
 CC -1- SURCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC PIR: A02113; OM0UB.
 DR SWISS-2DPAGE: P04217; HUMAN.
 DR MIM: 138670; .
 DR InterPro: IPRO01006; 1g_MHC.
 DR InterPro: IPRO01350; 1g_C2.
 DR InterPro: IPRO01600; 1g_LIKE.
 DR Pfam: PF00047; 1g; 4; .
 DR SMART: SMO0408; ICG2; 1.
 DR SMART: SMO0410; IC_LIKE; 4.
 DR Immunoglobulin domain: Glycoprotein: Plasma.
 KM DOMAIN 1 92 16-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN 1 93 185
 FT DOMAIN 1 94 206
 FT DOMAIN 1 95 228
 FT DOMAIN 1 96 250
 FT DOMAIN 1 97 272
 FT DOMAIN 1 98 294
 FT DOMAIN 1 99 316
 FT DOMAIN 1 100 338
 FT DISULFID 28 72
 FT DISULFID 402 449
 FT DISULFID 118 161
 FT DISULFID 211 258
 FT DISULFID 314 353
 FT DISULFID 354 393
 FT CARBOHYD 158 158
 FT CARBOHYD 342 342
 FT CARBOHYD 350 350
 FT CARBOHYD 350 350
 SEQUENCE 474 AA; 51940 MW; DA6CA53472A920EA CRC64;

Query Match 19 28; Score 325; DB 1; Length 474;
 Percent Similarity 71.7%; Expect 4.22E-177
 Matches 71; Conservative 3; Mismatches 17; Indels 8; Gaps 3;

OK NCBI_TaxID:11701;
 RN SEQUENCE FROM N.A.
 RA Mottelson B.R., Bain B.H., Shaw G.M., McNeely P.D., Modrov S.,
 RA Mottelson B.R., Bain B.H., Shaw G.M., McNeely P.D., Modrov S.,
 RA Wong-Staal P.; E.S., Forbes N.P., Josephs S.F., Gallo R.C.,
 RL Submitted (XXX-1987) to the HIV data bank.
 CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the University of Geneva. The content of this entry is the property of
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M17451; AAA5054.1; .
 DR HIV; M17451; 17SEP; Viral_Infect.
 DR Pfam: PF00559; VIF; 1
 DR PRINTS: PR00349; VIRIONINFECT.
 DR PRODOM: P000063; Viral_Infect; 1.
 DR AIDS.
 SEQUENCE 192 AA; 22687 MW; F005E0A621A5C6 CRC64;
 SO

Query Match 11 24; Score 74; DB 1; Length 192;
 Percent Similarity 22.0%; Expect 0.69;
 Matches 22; Conservative 14; Mismatches 40; Indels 24; Gaps 3;

OY 8 LLLNGV-----TWGPTVEALPIETVQXSLMAESRHLKTLGQCDADVPGPQDSRLPAV 61
 DB 8 MIVQGVDRERLRITKMSLVKHHNYISRKAKGMYRHHYESTIRPISSEVHIIPGDERLVIT 67
 OY 62 QEMGQEPHIDSAIKHCPITLTGQCYKRCGSLGTOW 101
 DB 68 TWG-----LHTGERD--WHGQCVSIEM 89

RESULT 3
 MW02_CHICK STANDARD: 663 AA.
 ID Q90611
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 20-AUG-2001 (rel. 40, Last annotation update)
 DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA
 DE GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).
 DR GELATINASE
 DR MW02; gallus (chicken)
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID:9031;
 RN SEQUENCE FROM N.A.
 RA Smith R.E., Smith R.E.; PubMed:8010954;
 RA MEDLINE:9320397; PubMed:8010954;
 RA Aimes R.T., French D.L., Quigley J.P.;
 RA "Cloning of a 72 kDa matrix metallopeptidase (gelatinase) from
 RA chicken embryo fibroblasts using gene family PCR: expression of the
 RA gelatinase increases upon malignant transformation.";
 RA Biochem. J. 300:729-736(1994).
 RL
 RL SEQUENCE OF 27-41 AND 107-122
 RA MEDLINE:9161603; PubMed:8484240;
 RA Chen J.-M., Aimes R.T., Ward G.R., Youngleat G.L., Quigley J.P.;
 RA "Isolation and characterization of a 70-kDa metalloproteinase
 RA (gelatinase) that is elevated in Rous sarcoma virus-transformed
 RA chicken embryo fibroblasts.";
 RA J. Biol. Chem. 266:5113-5121(1991).
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES

[illegible]

Query Match 10.7% Score 71: DB 1: Length 404;
 Best Local Similarity 21.5%: Pred. No. 3:4: Mismatches 28: Conservative 9: Mismatches 37: Indels 36: Gaps 5:

CC 4 LVFLLIMVYINQPVYEA-----AIFETQXSLMAWSRHKLTGDCDAVDP 51
 DB 12 LITLLIMVPGGEVYNAKAVITLDPKPMVIFPKENYITMCSGPHL----- 58
 CC 52 PQGQSLPAVDG---GAGPEVILSPAIKQIPLFDGQKTRRGSLG 98
 DB 59 -PCGSS-----TQWFINSTAVQISPVSYIFE---ASNDQSEKRCQICGS 100

RESULT 5
 VIF_HIVJR STANDARD: PRT: 192 AA.
 ID VIF_HIVJR
 AC 02087: 1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DE 01-JUL-1993 (Rel. 26, Last annotation update)
 CC VIRION INFECTIVITY FACTOR (SOR PROTEIN).
 GN Human immunodeficiency virus type 1 (HIV-1).
 CC Viruses: Retroid Viruses: Retroviridae: Lentivirus.
 CC F01_TaxID:11889;
 RA KEYMANI S., Chen I.S.Y.;
 RP SEQUENCE FROM N.A.
 DR InterPro: IPR00475: Viral_Infect.
 CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M36429; AA03746.1; -
 DR HIV: M36429; VIFSJRSP: Viral_Infect.
 DR InterPro: IPR00475: Viral_Infect.
 DR Pfam: PF00559: VIF: 1.
 DR PRINTS: PR00349: VIRIONINFECT.
 DR PRODOM: PD00065: Viral_Infect. 1.
 SN AIDS
 CC SEQUENCE 192 AA: 22443 MW: 82816986C100A6F0 CSMC4:

Query Match 10.4% Score 69: DB 1: Length 192;
 Best Local Similarity 21.0%: Pred. No. 2:3: Mismatches 21: Conservative 15: Mismatches 40: Indels 24: Gaps 3:

CC 8 LILKSVY-----TQGVTEALIEGTQXSLMAWSESHKLTGDCDAVDPGQDQSLPAV 61
 DB 8 LILKSVY-----TQGVTEALIEGTQXSLMAWSESHKLTGDCDAVDPGQDQSLPAV 61
 CC 62 QEMGAGPEVILSPAIKQIPLFDGQKTRRGSLG 101
 DB 62 QEMGAGPEVILSPAIKQIPLFDGQKTRRGSLG 101
 CC 68 TYWG-----LITGERD---WHLSCQVSMKW 89
 DB 68 TYWG-----LITGERD---WHLSCQVSMKW 89

RESULT 6
 OSJL_ARAVH STANDARD: PRT: 244 AA.
 ID OSJL_ARAVH
 AC P50700:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 20-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 CC HLA-B OR HLA-B
 CC CHAIN PRECURSOR.
 CC HLA-B OR HLA-B
 CC Homo sapiens (Human).
 CC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
 CC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 CC NCBI_TaxID:9606;

Arabisopsis thaliana (Mouse-ear cress).
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 CC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 CC eurosidus 11: Brassicales: Brassicaceae: Arabidopsis.
 CC NCBI_TaxID:3702;
 CC SEQUENCE FROM N.A.
 CC STRAIN:CV. COLUMBIA; TISSUE: Leaf;
 CC MEDLINE:97354294; PubMed:92105083;
 CC Gapped N., Dlogon T., Grepin H., Simon P.;
 CC Isolation and characterization of a cDNA clone encoding an
 CC protein of 1556 (1997).
 CC NCBI_TaxID:3702;
 CC -1- SIMILARITY: RELATES TO THE THAUMATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X89008; CA61411.1; -
 DR HSP: 825871: 140N.
 DR InterPro: IPR001938: Thaumatin.
 DR Pfam: PF00314: Thaumatin. 1.
 DR PRODOM: PD001321: THAUMATIN. 1.
 DR SMART: SM00205: THN: 1.
 DR PROSITE: PS00316: THAUMATIN: 1.
 CC SIGNAL
 CC POTENTIAL.
 CC SIGNAL 23 22
 CC OSNOTIN-LIKE PROTEIN OSM34.
 CC SIGNAL 23 22
 CC BY SIMILARITY.
 CC DISULFID 71 82
 CC BY SIMILARITY.
 CC DISULFID 87 93
 CC BY SIMILARITY.
 CC DISULFID 138 212
 CC BY SIMILARITY.
 CC DISULFID 143 195
 CC BY SIMILARITY.
 CC DISULFID 151 161
 CC BY SIMILARITY.
 CC DISULFID 165 174
 CC BY SIMILARITY.
 CC DISULFID 175 182
 CC BY SIMILARITY.
 CC SEQUENCE 244 AA: 26603 MW: 576839569F2865D0 CSMC4:

Query Match 10.4% Score 68.5: DB 1: Length 244;
 Best Local Similarity 32.5%: Pred. No. 3:5: Mismatches 25: Conservative 4: Mismatches 27: Indels 21: Gaps 4:

CC 42 LGGCDADV---PQPGQSLPAVDGQKTRRGSLG-----DTQCHT 92
 DB 28 LGGCDADV---PQPGQSLPAVDGQKTRRGSLG-----DTQCHT 92
 CC 93 CRSG-----LSTGKXQ 103
 DB 93 CRSG-----LSTGKXQ 103
 CC 82 CQYDQSGGGLQCTGCGO 98
 DB 82 CQYDQSGGGLQCTGCGO 98

RESULT 7
 IB55_HUMAN STANDARD: PRT: 362 AA.
 ID IB55_HUMAN
 AC P30492:
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 CC HLA-B OR HLA-B
 CC CHAIN PRECURSOR.
 CC HLA-B OR HLA-B
 CC Homo sapiens (Human).
 CC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
 CC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 CC NCBI_TaxID:9606;

SEQUENCE FROM N.A.
 MEDLINE:92148136; PubMed:1737933;
 RA Hildebrand W.H., Madrigal J.A., Little A.-M., Parham P.;
 RA HLA-B*22: a family of molecules with identity to HLA-B7 in the alpha
 1-helix.; 148-1155-1162(1992).
 CC -1- FUNCTION INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC -1- THE IMMUNE SYSTEM
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 MICROGLOBULIN).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its contents are not
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: 477774; AA03086.1; .
 CC MIM: 142830; .
 CC DR InterPro: IPR003006; 19_MHC.
 CC DR InterPro: IPR003597; 19_CL.
 CC DR InterPro: IPR001039; MHC_I.
 CC Pfam: PF00047; 19; 1.
 CC PRODOM: PD000050; MHC_I; 1.
 CC PROSITE: PS00250; 1001; MHC_I.
 CC MHC I; Transmembrane; Glycoprotein; Signal.
 CC SIGNAL 1 24
 CC CHAIN 25 362 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
 CC DONAN 25 114 BW-54(BW-22) B*5401 ALPHA CHAIN.
 CC DONAN 115 206 EXTRACELLULAR ALPHA-1.
 CC DONAN 115 206 EXTRACELLULAR ALPHA-2.
 CC DONAN 259 308 EXTRACELLULAR ALPHA-3.
 CC TRANSMEM 309 332 CONNECTING PEPTIDE.
 CC DONAN 333 362 CYTOPLASMIC TAIL.
 CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 CC DISULFID 125 188 BY SIMILARITY.
 CC FT 227 283
 CC SEQUENCE 362 AA; 40380 MW; 11730024578515D CRC64;

 CC Query Match 10.1%; Score 67; DB 1; Length 362;
 CC Best Local Similarity 24.2%; Pred. No. 7.9;
 CC Matches 32; Conservative 14; Mismatches 64; Indels 22; Gaps 5;

 CC 6 VPLLIVGV---TNGPVDAIFVETFGXSIANSEHXIKTLG---QCDADVGP 52
 CC DB 9 LULLSLPALATLTERVMSGSHRTY-TAKNSNGKGEPRVLAICYDGTGCFVPSDANSP 67
 CC OY 53 RPSDSLPAVQKMAQEPVHLSDPILHOFILFDGKRGKSGXGKXSLLEETG 112
 CC DB 68 RGEPAFAVVEQEG--PEVYDNTQIYKAQADRESLNLNR-----GYNSQESGSHW 119
 CC OY 113 PXYLACSLADG 124
 CC DB 120 QTYGCDLQDFPG 131

 CC RESULT 8
 CC Y4PA_RH1SN STANDARD; PRT; 609 AA.
 CC AC P55610;
 CC DT 01-NOV-1997 (Rel. 35; Created)
 CC DT 01-NOV-1997 (Rel. 35; Last sequence update)
 CC DT 01-NOV-1997 (Rel. 35; Last annotation update)
 CC DE PUTATIVE TRANSCRIPTIONAL REGULATORY PROTEIN Y4PA.
 CC GN Y4PA.
 CC Rhizobium sp. (strain NGR234).

Plasmid sym pNCR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OC NCBI_TaxID=394;

 CC SEQUENCE FROM N.A.
 CC MEDLINE:92193555; PubMed:9154424;
 CC RA Pether X., Pelly R., Bairoch A., Broughton W.J., Rosenthal A.;
 CC RA Pether X.;
 CC RL Nature 387:394-401(1997).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR THAT ACTS IN
 CC CONDUCTION WITH SIGMA-54.
 CC -1- INTERACT: THE ENHANCER REGION CONTAINS A SIGMA-54 FACTOR
 CC -1- SIMILARITY: IN THE N-TERMINAL TO Y4OF AND Y4OV.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its contents are not
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL: A600089; AAB9181.1; .
 CC DR InterPro: IPR002197; RTH_FIS.
 CC DR InterPro: IPR002078; S1954_Interact.
 CC Pfam: PF00138; 2; Sigma54; 2.
 CC PROSITE: PS00676; SIGMA54_INTERACT_1; FALSE_NEG.
 CC PROSITE: PS00688; SIGMA54_INTERACT_2; FALSE_NEG.
 CC PROSITE: PS00688; SIGMA54_INTERACT_3; FALSE_NEG.
 CC PROSITE: PS00688; SIGMA54_INTERACT_4; 1.
 CC Hypothetical protein; Transcription regulation; DNA-binding;
 CC ATP-binding; Plasmid.
 CC DONAN 313 533 SIGMA-54 FACTOR INTERACTION (POTENTIAL).
 CC FT NR_BIND 325 604 ATP (POTENTIAL) (SIMILARITY).
 CC FT 574 574
 CC SEQUENCE 609 AA; 66700 MW; 8CC727E67D508F36 CRC64;

 CC Query Match 10.1%; Score 67; DB 1; Length 609;
 CC Best Local Similarity 30.3%; Pred. No. 14;
 CC Matches 23; Conservative 9; Mismatches 18; Indels 26; Gaps 5;

 CC 32 MAFSHYKTYGCOADVGP-RGDSRLPAVQKMAQEPVHLSDPILHOFILFDGKRGKSGXGKXSLLEETG 90
 CC DB 539 WIKSEH-LPPLGDANNDAPPHIDRER---EW-----IL-DALOR 573
 CC OY 91 YKNSGSLGKXSLK 106
 CC DB 574 HRFKGEARLYLGS 589

 CC RESULT 9
 CC HEM1_HELMO STANDARD; PRT; 443 AA.
 CC AC O9ZG6;
 CC DT 30-MAY-2000 (Rel. 39; Created)
 CC DT 20-MAY-2000 (Rel. 39; Last sequence update)
 CC DT 20-MAY-2001 (Rel. 40; Last annotation update)
 CC GN HEM1-TRNA REDUCTASE (Ec 1.2.1.-) (GUTR).
 CC OS Hellobacillus mobilis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Hellobacterium group; Hellobacillus.
 CC NCBI_TaxID=28064;
 CC SEQUENCE FROM N.A.
 CC MEDLINE:9601357;
 CC Xiong J., Inoue K., Bauer C.E.;
 CC "Tracking molecular evolution of photosynthesis by characterization of
 CC a major photosynthesis gene cluster from Hellobacillus mobilis.";

[illegible]

```

DR HSP90: P00808; IMBL.
DR TIGR: MT2128; -.
DR TuberculinS: RA20569C; -.
DR InterPro: IPRO01456; Beta_Lactam.
DR InterPro: IPR000871; Beta_Lactam_A.
DR Pfam: PF00141; beta-lactamase; 1.
DR PROSITE: PS00013; BETA-LACTAMASE.
DR PROSITE: PS00013; BROKRACTOPROTECTIN.
DR PROSITE: PS00146; BETA-LACTAMASE A; 1.
KW Hydrolase; Antibiotic resistance; Membrane; Lipoprotein; Signal;
KW Complete proteome.
FT SIGNAL          1      23
FT CHAIN           24     307
FT LIPID            84      84
FT ACET_SITE       84      84
FT BINDING         250     252
FT SDOQUENCE       307 AA; 44SCZADMSDFY413 CRG64;
SQ
Query Match          9.8%; Score 65; DB 1; Length 307;
Best local Similarity 27.9%; Pred. No. 11;
Matches 36; Conservative 17; Mismatches 44; Indels 32; Gaps 7;
Oy   12 GVTMGPTAEALAFYE---TKVSL-----VAESENHXIKTIG--CCDADV-----50
Db   128 EMTIISGLCDLARSISQGNANLLADLGPGCGTAATFYGISRGVYSRLAEEPELMIR 187
Oy   51 CGPGRSR---LPANVGEMSGDEPHMLDSPIKNTDLNGF---DPGGARREGSGSTGXNO 103
Db   188 DPCHSDERTPTTHNALVALDQLGVALLPKPCBAALLTDWMAKNRTCAKRITACGPADM---245
Oy   104 LSKRLLELNG 312
Db   1:::11

```

```

DB 246 - -KVIDRTG 252

RESULT 15
ID V130 VZVD STANDARD; PRT: 483 AA.
AC P03276:
RT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE CASP12 ASSEMBLY AND DNA MATURATION PROTEIN (CASP12 PROTEIN VP19C).
GN 20.
CS Casp12-zoster virus (strain Dumas) (VZV).
CC Virus; deDNA; herpesviridae.
OC Alphaherpesvirinae: Varicellovirinae.
OK NCBI_TaxID:10336;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson A.J., Scott J.E.,
MEDLINE:86306657; PubMed:3018124;
RT The zoster virus (VZV) is a member of the gamma-herpesvirinae subfamily.
CC -1- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CASPIDS ARE
EMBEDDED. BINDS DNA (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CASP12 PROTEIN VP19C.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X01470; CA227903.1; -
DR EMBL: X01470; CA227903.1; -
DR CASP12 ASSEMBLY AND DNA MATURATION PROTEIN (CASP12 PROTEIN VP19C).
OK SEQUENCE 483 AA: 53971 MW: 4584C73D689F9A1 CRC64:

Query Match 9.8%; Score 65; DB 1; Length 483;
Best Local Similarity 31.4%; Pred. No. 18;
Matches 16; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

OY 56 SRPAPGNGAGPVPALDPAIKIOFLTGTGGVYVCSGSGXSLTGMQK 106
DB 297 ANIPACVFMDKDLHLSADCKHVELFVYVGRORRQECVRLHLSQSLME 347

RESULT 16
AC P98177: 013720: 043821; PRT: 504 AA.
RT 01-OCT-1996 (Rel. 34, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE PUTATIVE FORK HEAD DOMAIN TRANSCRIPTION FACTOR AFX1.
GN MULT7 OR AFX1 OR AFX.
CC Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OK NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TITSIDE-Blood;
MEDLINE:9510180; PubMed:9341872;
RT AFX1 and p57 are two proteins that are highly homologous in their
CC AFX1 and p57 are two proteins that are highly homologous in their
CC candidate genes of X-linked dysostosis parkinsonism.*;
RN Hum. Gene. 100:569-572(1997).
RP SEQUENCE FROM N.A.
RA MEDLINE:97163401; PubMed:9010221;

```

```

RA Borthardt A., Pepp R., Haas O.A., Lets T., Harbott J., Kreuder J.,
RT Hammermann J., Henn T., Lampert F.,
CC *Cloning and characterization of AFX, the gene that fuses to MLL in
RT acute leukemias with a t(4;11)(q31;q23)*.
RL Oncogene 14:195-202(1997).
RN [1]
RP CHROMOSOMAL TRANSLOCATION.
RX TISSUE=bone marrow;
MEDLINE:95118921; PubMed:7529552;
RT Parry P., Mel Y., Evans G.;
CC *Cloning and characterization of the t(4;11) breakpoint from a
RT leukemic cell line identify a new member of the forkhead gene
CC family.*;
RN Hum. Gene. 100:569-572(1997).
CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
MUSCLE, KIDNEY AND PANCREAS.
CC -1- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY A CHROMOSOMAL
TRANSLOCATION T(4;11)(Q13;Q23) THAT INVOLVES MULT7 AND MLL/HRX.
CC THE RESULT IS A ROGUE ACTIVATOR PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: Y11284; CA273156.1; -
DR EMBL: Y11285; CA273156.1; JOINED.
DR EMBL: Y11285; CA273156.1; JOINED.
DR EMBL: X93956; CA63819.1; -
DR MIM: 300003; A0001765; Fork head.
DR PRT: PR00250; Fork head.
DR PRT: PR00053; FORKHEAD.
DR SMART: SM00339; FH. 1.
DR PROSITE: PS00657; FORK-HEAD_1; FALSE NEG.
DR PROSITE: PS00658; FORK-HEAD_2; 1.
DR PROSITE: PS50039; FORK-HEAD_3; 1.
RN NM_000001; MULT7; Fork head gene.
CC Transcription regulation; DNA-binding; Nuclear protein;
FT DNA BIND 99 187
FT CONFLICT 1 33
FT CONFLICT 73 73 P -> S (IN REF. 2).
FT CONFLICT 78 78 A -> G (IN REF. 2).
FT CONFLICT 108 108 L -> F (IN REF. 2).
FT CONFLICT 121 121 L -> I (IN REF. 2).
OK SEQUENCE 504 AA: 53470 MW: 3776754VFD3286C4 CRC64:

Query Match 9.8%; Score 65; DB 1; Length 504;
Best Local Similarity 28.0%; Pred. No. 19;
Matches 33; Conservative 16; Mismatches 45; Indels 24; Gaps 7;

OY 12 CYNMGPV--TAAATFETGKSLA---SESRHIXKTLGGDGDVPPGSGSLPRANGEM 64
DB 332 GVT--CPALHYSSSFSPARCPPLASGCCGSSQALTEAL--LTSDIPPPADVLMTQV--- 385
OY 65 GAGPVPALDPAIKIOFLTGTGGVYVCSGSGXSLTGMQKSLLETCGPVYACSLAL 122
DB 386 ---DPLLSQAFTL---LLGLGLPS---SSKRLAVGCLCPRLNAPGSPSLVFTSLW 432

RESULT 17
AC P39650: STANDARD; PRT: 258 AA.
RT RSFL_BACSU

```


OK NCBI_TaxID=9606;
 RA SEQUENCE FROM N.A. PubMed:1384166;
 RA MEDLINE=9301775; Hildebrand W.H., Ward F.E., Marsh S.G.,
 RA Zemmour J., Gumpertz J.E., Hildebrand W.H., Ward F.E., Marsh S.G.,
 RA Williams R.C., Parham P.;
 RA "The molecular basis for reactivity of anti-Cw1 and anti-Cw3
 RA allantoisera with HLA-B*6 haplotypes";
 RA Tissue Antigens 39:249-257(1992).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send an email to license@sb.stb.ch).
 CC EMBL: M64174; AA595688.1;
 CC HSSP: P30460; JAGB.
 DR HMM: 142840;
 DR Interpro: IPR003006; Iq_MHC.
 DR Interpro: IPR003597; Iq_C1.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR ProDom: PD000050; MHC.I.1.
 DR SMART: SM00407; ICG1.1.
 DR PROSITE: PS00290; IG_MHC.1.
 DR MHC.I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 366
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 333
 FT TRANSMEM 334 366
 FT DISULFD 325 188
 FT DISULFD 227 283
 FT CARBOHYD 110 110
 SEQUENCE 366 AA: 40772 MW: 24844D1399A0486A CRC64;
 Query Match 9.5%; Score 63; DB 1; Length 366;
 Best Local Similarity 24.8%; Pred. No. 21;
 Matches 34; Conservative 14; Mismatches 67; Indels 22; Gaps 5;
 Oy 1 MSALVPELLMGV-----TWGVTPEAFAFETQXSLWASERKATG-----OCDA 47
 Db 4 MARETLILLSGALALFETWASHSRFFY-TASRSGSGPRFAVAGYDDOTFQVDS 62
 Oy 48 DVGFPDSSRLPVDGMAOEIVHDSFAIKIQLPLGTGTGGRSGSLTGKXQLSKL 107
 Db 63 DAASPRGEPRAPVWDEDS--PEYWRQETQKYKQAOETRVSLRLNR-----GYWQSEA 114
 Oy 108 LETGPVYACSLALDG 124
 Db 115 GSHTLGRMYGCDLGPDG 131
 RESULT 23
 ID 1C13_HUMAN STANDARD: PRI: 366 AA.
 AC P30506;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, CW-8 CW*0802 ALPHA CHAIN
 DE PROCDROR (CMB.2).
 GN HLA-C OR HLA-C.
 OS Homo sapiens (human)
 OS Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RA SEQUENCE FROM N.A. PubMed:1384166;
 RA MEDLINE=9301775; Hildebrand W.H., Ward F.E., Marsh S.G.,
 RA Zemmour J., Gumpertz J.E., Hildebrand W.H., Ward F.E., Marsh S.G.,
 RA Williams R.C., Parham P.;
 RA "The molecular basis for reactivity of anti-Cw1 and anti-Cw3
 RA allantoisera with HLA-B*6 haplotypes";
 RA Tissue Antigens 39:249-257(1992).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb.stb.ch/annouce/
 CC or send an email to license@sb.stb.ch).
 CC EMBL: M64173; AA595687.1;
 CC HSSP: P30460; JAGB.
 DR HMM: 142840;
 DR Interpro: IPR003006; Iq_MHC.
 DR Interpro: IPR003597; Iq_C1.
 DR Interpro: IPR001039; MHC.I.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR Pfam: PF00047; Iq_1.
 DR ProDom: PD000050; MHC.I.1.
 DR SMART: SM00407; ICG1.1.
 DR PROSITE: PS00290; IG_MHC.1.
 DR MHC.I; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 366
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 298
 FT DOMAIN 299 308
 FT TRANSMEM 309 333
 FT TRANSMEM 334 366
 FT DISULFD 325 188
 FT DISULFD 227 283
 FT CARBOHYD 110 110
 SEQUENCE 366 AA: 40871 MW: D343B09436B8132C CRC64;
 Query Match 9.5%; Score 63; DB 1; Length 366;
 Best Local Similarity 24.8%; Pred. No. 21;
 Matches 34; Conservative 14; Mismatches 67; Indels 22; Gaps 5;
 Oy 1 MSALVPELLMGV-----TWGVTPEAFAFETQXSLWASERKATG-----OCDA 47
 Db 4 MARETLILLSGALALFETWASHSRFFY-TASRSGSGPRFAVAGYDDOTFQVDS 62
 Oy 48 DVGFPDSSRLPVDGMAOEIVHDSFAIKIQLPLGTGTGGRSGSLTGKXQLSKL 107
 Db 63 DAASPRGEPRAPVWDEDS--PEYWRQETQKYKQAOETRVSLRLNR-----GYWQSEA 114
 Oy 108 LETGPVYACSLALDG 124
 Db 115 GSHTLGRMYGCDLGPDG 131
 RESULT 24

Db	353	YLLSTYAMGACGATGTSQCAMKPLCSASHAFQYONFLAMALSTDRKLPKSTAKADMAA	412
Qy	42	-----LQGDADVGEFPCDSRL	58
Db	413	SNORLEPTMILQNSGCIACGATNSMCAVADPACTGTTCGATGAAYVDPSNM	472
Qy	59	PAVQSGKQPVHDSFAMHGFLLTTPDQGRGRCRGLSTGKXKMSLLETPTQALAC	118
Db	473	FGQMGAGVQAVDEL-----YASQNAQAK-----KLDKAVPMVVA--	508
Qy	119	SLAIDGAS	126
Db	509	NITDGAS	516
RESULT 28			
CPZB_HUMAN		STANDARD:	PRT: 508 AA.
1C	CPZB_HUMAN		
ID	CP15529		
NC	15-JUL-1998 (Ref. 36, Created)		
DT	15-JUL-1998 (Ref. 37, Sequence update)		
DT	20-MIG-2001 (Ref. 40, Last annotation update)		
DE	25-HYDROXYVITAMIN D-1 ALPHA-HYDROXYLASE, MITOCHONDRIAL PRECURSOR		
DE	(EC 1.14.-.-) (25-OH-D-1-ALPHA-HYDROXYLASE) (25-HYDROXYVITAMIN D3 1-		
DE	ALPHA-HYDROXYLASE) (VVD) 1A HYDROXYLASE) (P450C1 ALPHA) (P450VD1-		
DE	ALPHA)		
GN	CPZB2H1 OR CPZB7B OR CPZB1ALPHA.		
GN	Bikayeva M, et al. Chordata: Chelonia; Vertebrata: Euteleostomi:		
GN	Bikayeva M, et al. Chordata: Chelonia; Vertebrata: Euteleostomi:		
GN	Memorial: Euteleostomi: Primates: Catarrhini: Hominae: Homo.		
GN	NCBI_TaxID=9606;		
RP	11		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE:98083075; PubMed:9428739;		
RP	F.O.C.R., Portale A.P., Miller W.L.,		
RP	hydroxylase P450c1alpha *, human gene for the vitamin D 1alpha-		
RP	DNA Cell Biol. 16:1499-1507(1997).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE: kidney;		
RP	MEDLINE:98008873; PubMed:9344864;		
RP	Komatsu T., Yoshida T., Makino S., Shinki T., Anazawa H.,		
RP	Molecular cloning of cDNA and genomic DNA for human		
RP	25-hydroxyvitamin D3 1 alpha-hydroxylase *;		
RP	Biochem. Biophys. Res. Commun. 239:527-533(1997).		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE:98075882; PubMed:9415400;		
RP	F.O.C.R., Lin D., Zhang T.H., Biale D.D., Shackleton C.H., Miller W.L.,		
RP	Portale A.P.,		
RP	cloning of human 25-hydroxyvitamin D-1 alpha-hydroxylase and		
RP	mutations causing vitamin D-dependent rickets type 1.*;		
RP	Mol. Biotechnol. 11:1961-1970(1997).		
RP	1- FUNCTION: CATALYZES THE CONVERSION OF 25-HYDROXYVITAMIN D3		
RP	(25(OH)D) TO 1-ALPHA,25-DIHYDROXYVITAMIN D3 (1,25(OH)2D) PLAYS AN		
RP	IMPORTANT ROLE IN CRITICAL ROLE IN NORMAL BONE GROWTH, CALCIUM		
RP	1- PRIMARY SECOND STEP IN THE CONVERSION OF VITAMIN D(3) INTO THE		
RP	ACTIVE FORM (1-ALPHA,25-DIHYDROXYVITAMIN D(3)).		
RP	1- SUBCELLULAR LOCATION: MITOCHONDRIAL.		
RP	1- TISSUE SPECIFICITY: KIDNEY.		
RP	1- DISEASE: DEFECTS IN CPZB7B ARE A CAUSE OF VITAMIN D-DEPENDENT		
RP	RICKETS TYPE 1 (VDDR-1), A DISEASE CHARACTERIZED BY MUSCLE		
RP	WEAKNESS AND RICKETS.		
RP	1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
CC	-----		
CC	THIS SWISS-PORT ENTRY IS COPYRIGHT, IT IS PRODUCED THROUGH A COLLABORATION		
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -		
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS		
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY		

[illegible]

HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, CW-2, CW-0202 ALPHA CHAIN
 DE PRECURSOR (CW2.2).
 DE HLA-C OR HLA-C.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OS Euteleostomi; Actinopterygii; Cyprinodontiformes; Zebrafish;
 OS Mammalia; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 [11]
 NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89235215; PubMed=2715640;
 RA Parham P., Lawlor D.A., Lomen C.E., Ennis P.D.;
 RA Diversity and diversification of HLA-A,B,C alleles.*
 RT J. Immunol. 142:3535-3540(1989).
 CC This sequence was derived from the PRESENTATION OF FOREIGN ANTIGENS TO
 CC THE IMMUNE SYSTEM.
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. Bioinformatics institutions are encouraged to use the data for commercial
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC
 CC EMBL: M24030; AAAS5971.1;
 CC
 CC HSP: P30685; IAIN.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003597; I9_C1.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; I9: 1.
 DR Pfam: PF00129; MHC_I: 1.
 DR PROSITE: PS00250; MHC_I: 1.
 DR PROSITE: PS00250; I9_C1: 1.
 DR PROSITE: SM00407; I9_C1: 1.
 DR PROSITE: SM00407; I9_C1: 1.
 DR MHC_I: Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 366
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 288
 FT DOMAIN 289 308
 FT TRANSMEM 309 333
 FT DOMAIN 334 366
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 SQ SEQUENCE 366 AA: 40994 MW: 64AD7B8C3BD5D6 CRC64;
 N-TERMINAL (GLNMG...) (BY SIMILARITY).
 Query Match 9.4%; Score 62; DB 1; Length 366;
 Best Local Similarity 24.1%; Pred. No. 27;
 Matches 33; Conservative 17; Mismatches 65; Indels 22; Gaps 5;
 1 MSMLVETLLAGV-----TWGPEVTEALFETVQXSLAESHKATLTC-----QCDA 47
 4 MAFPTLLTSLGALVETWAGSHSRKRY-TVASRSGSGPRTAVGVDDTPVFRDS 62
 48 DVCSPGDSRLPVDQMGAEFVHLSDPAIKIQLPLGCTGCGYRCSGLSTGCKXSL 107
 63 DAAISRGCEPAPVAVROEC--PEYWMREFQKVRQADOTGVNRLKRLR-----GYNOSEA 114
 108 LEITGPVLVSLADLG 124
 115 GSHTLQKNCQGLADPG 131
 RESULT 32
 1C1S_HUMAN STANDARD; PRT; 366 AA.
 ID 1C1S_HUMAN

HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, CW-1201 ALPHA CHAIN PRECURSOR
 DE 01-APR-1993 (Rel. 25, Created)
 DE 01-APR-1993 (Rel. 25, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 GN HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, CW-1201 ALPHA CHAIN PRECURSOR
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OS Mammalia; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 [11]
 NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89351040; PubMed=2843451;
 RA Takeda H., Inoko H., Ando A., Matsunaka M., Matsunaka B., Tsuji K.,
 RT "Cloning and analysis of HLA class I cDNA encoding a new HLA-C
 RT *specificity Cx52.*"
 CC Immunogenetics 28:265-270(1988).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC
 CC EMBL: M21963; AAAS947.1;
 CC
 CC HSP: P30685; IAIN.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003597; I9_C1.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; I9: 1.
 DR Pfam: PF00129; MHC_I: 1.
 DR PROSITE: PS00250; MHC_I: 1.
 DR PROSITE: SM00407; I9_C1: 1.
 DR PROSITE: SM00407; I9_C1: 1.
 DR PROSITE: SM00290; I9_MHC: 1.
 DR MHC_I: Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 366
 FT DOMAIN 25 114
 FT DOMAIN 115 206
 FT DOMAIN 207 288
 FT DOMAIN 289 308
 FT TRANSMEM 309 333
 FT DOMAIN 334 366
 FT DISULFID 125 188
 FT DISULFID 227 283
 FT CARBOHYD 110 110
 SQ SEQUENCE 366 AA: 40851 MW: EADVAISLTVN546 CRC64;
 N-TERMINAL (GLNMG...) (BY SIMILARITY).
 Query Match 9.4%; Score 62; DB 1; Length 366;
 Best Local Similarity 24.1%; Pred. No. 27;
 Matches 33; Conservative 15; Mismatches 67; Indels 22; Gaps 5;
 1 MSMLVETLLAGV-----TWGPEVTEALFETVQXSLAESHKATLTC-----QCDA 47
 4 MAFPTLLTSLGALVETWAGSHSRKRY-TVASRSGSGPRTAVGVDDTPVFRDS 62
 48 DVCSPGDSRLPVDQMGAEFVHLSDPAIKIQLPLGCTGCGYRCSGLSTGCKXSL 107
 63 DAAISRGCEPAPVAVROEC--PEYWMREFQKVRQADOTGVNRLKRLR-----GYNOSEA 114
 108 LEITGPVLVSLADLG 124
 115 GSHTLQKNCQGLADPG 131

```

DB 63 DAASRGEPRAPWVEDEG---PEYMWRETKYKROADRVSLRLRLR-----GYNOSSEA 114
OY 108 LBTGPVKVLAASLADG 124
DB 115 GSRHLOANNICDGLADPG 131

RESULT 33
CG95_HUMAN STANDARD: PRT: 620 AA.
ID CG95_HUMAN 1997 (rel. 35. Created)
RT 0-NOV-1997 (rel. 35. Last sequence update)
DT 20-AUG-2001 (rel. 40. Last annotation update)
DE GOLG1N-95.
GN GOLG1N-95.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC 111 TaxID:9606;
RP TISSUE: LIVER;
RP SEQUENCE FROM N.A.
RX MEDLINE:9301617; PubMed:8315394;
RX Feltzler M.J., Hamel J.C., Ochs R.L., Chan E.K.L.;
RX "Molecular characterization of two human autoantigens: unique cDNAs
RX encoding 95- and 160-kD proteins of a putative family in the golgi
RX 3' exp. Mod. 178:49-62(1993).
-1- FUNCTION: GOLGI AUTO-ANTIGEN; MAY HAVE A FUNCTION IN THE
-1- PROCESSING AND TRANSPORT OF PROTEINS THROUGH THE GOLGI.
-1- DOMAIN: EXTENDED ROD-LIKE PROTEIN WITH COILED-COIL DOMAINS.
-1- SIMILARITY: HIGH, TO RAT CIS-GOLGI MATRIX PROTEIN GM130.
-1- This SWISS-PROF entry is copyright. It is produced through a collaboration
-1- between the Swiss Institute of Bioinformatics and the EMBL outstation
-1- at the European Bioinformatics Institute. There are no restrictions on its
-1- use by non-profit institutions as long as its content is in no way
-1- modified and this statement is not removed. Usage by and for commercial
-1- entities requires a license agreement (See http://www.isb-sib.ch/announcements
-1- or send an email to license@isb-sib.ch).
DB EMBL: L06147; AAA35920.1; -.
DR MIM: 602580; -.
KM GOLGI stack: Coiled coil.
FT DOMAIN 5 336 COILED COIL (POTENTIAL).
FT DOMAIN 352 510 COILED COIL (POTENTIAL).
FT DOMAIN 57 73 POLY-PRO.
FT FT 32 333 POLY-GLU.
FT SEQUENCE 620 AA: 70477 MW: 54831 AND PMR42PF08 CRC64;
OY 76 EYEGQICAAEHRIKRELEIGLADQIQAOYQONEDLSRLNREGRRLLELRKAELKMEQ 135
DB 115 GSRHLOANNICDGLADPG 131

Query Match 9.4%; Score 62; DB 1; Length 620;
Best Local Similarity 36.2%; Pred. No. 50;
Matches 21; Conservative 2; Mismatches 19; Indels 16; Gaps 2;

26 EYQXSLWASERKAKLT-----GQCDADVPQGFQDSRIKPAVDE-----WCAQ 67
DB 115 GSRHLOANNICDGLADPG 131
OY 76 EYEGQICAAEHRIKRELEIGLADQIQAOYQONEDLSRLNREGRRLLELRKAELKMEQ 135
DB 115 GSRHLOANNICDGLADPG 131

RESULT 34
Y029_HUMAN STANDARD: PRT: 971 AA.
ID Y029_HUMAN 1998 (rel. 36. Created)
RT 15-JUL-1998 (rel. 36. Last sequence update)
DT 20-AUG-2001 (rel. 40. Last annotation update)
DE HYPOTHEMETICAL PROTEIN KIAA0029.
GN KIAA0029.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

```

OX NCBI_TaxID:9606;
RN 111;
RP SEQUENCE FROM N.A.
RX MEDLINE:96001387; PubMed:7584026;
RX "The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RX RT analysis of randomly sampled cDNA clones from human immature myeloid
RX cell line KC-1."
RT DNA Res. 1:27-35(1994).
-1- This SWISS-PROF entry is copyright. It is produced through a collaboration
-1- between the Swiss Institute of Bioinformatics and the EMBL outstation
-1- at the European Bioinformatics Institute. There are no restrictions on its
-1- use by non-profit institutions as long as its content is in no way
-1- modified and this statement is not removed. Usage by and for commercial
-1- entities requires a license agreement (See http://www.isb-sib.ch/announcements
-1- or send an email to license@isb-sib.ch).
DB EMBL: D21855; BA044878.1; -.
DR InterPro: IPR00374; R3H.
DR Pfam: PF01424; R3H; 1.
DR SMART: SM00393; R3H; 1.
DR Hypothetical protein.
FT DOMAIN 461 480 POLY-PRO.
FT SEQUENCE 971 AA: 107673 MW: D07684D368955108 CRC64;
OY 100 ---CMXOJLSKLELV 111
DB 848 LQNGEYVQKYLELT 862

Query Match 9.4%; Score 62; DB 1; Length 971;
Best Local Similarity 29.3%; Pred. No. 85;
Matches 22; Conservative 11; Mismatches 24; Indels 18; Gaps 4;

49 YVPPGDSPLAVQEMKADPEVHDSPIKQGLITGDT-----GRYKCSQST- 99
DB 794 VPG--QDMSRYPL-----GQPLQYNPVAVHGHIFNQDQPGSRIGNRGRKAKKASTD 847
OY 100 ---CMXOJLSKLELV 111
DB 848 LQNGEYVQKYLELT 862

RESULT 35
ACAL_ARATH STANDARD: PRT: 1020 AA.
ID ACAL_ARATH 20-AUG-2001 (rel. 40. Created)
RT 20-AUG-2001 (rel. 40. Last sequence update)
DT 20-AUG-2001 (rel. 40. Last sequence update)
DE CALCIUM-TRANSPORTING ATPASE 1 PLASMA MEMBRANE TYPE (EC 3.8.3.8)
DB ACAL OR PEAL OR AT1G27770 OR T265.23 OR F2BL5.1.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbiales; Malvales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID 3702;
RN 111;
RP PRELIMINARY SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN:CV. COLUMBIA;
RX MEDLINE:94052104; PubMed:8234257;
RX Huang L., Berkeiman T., Franklin A.E., Hoffman N.B.;
RX "The plasticization of a gene encoding a Ca(2+)-ATPase-like protein in
RX P. nelt. Acad. Sci. U.S.A. 90:10066-10070(1993).
RP REVISIONS:
RP Huang L., Berkeiman T., Franklin A.E., Hoffman N.E.;
RP Proc. Natl. Acad. Sci. U.S.A. 91:9664-9664(1994).
RP [3]
RP SEQUENCE FROM N.A.

```

[illegible]


```

FT CHAIN 174 276 32 KDA ENAMELIN.
FT CHAIN 515 665 25 KDA ENAMELIN.
FT CHAIN 670 7 34 KDA ENAMELIN.
FT MOD_RES 53 53 PHOSPHORYLATION.
FT MOD_RES 191 191 PHOSPHORYLATION.
FT MOD_RES 216 216 PHOSPHORYLATION.
FT MOD_RES 547 547 HYDROXYLATION.
FT CARBOHYD 245 245 N-LINKED (GLCNAc. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAc. . .).
FT CARBOHYD 291 291 N-LINKED (GLCNAc. . .).
FT CARBOHYD 462 462 N-LINKED (GLCNAc. . .).
FT CARBOHYD 929 929 N-LINKED (GLCNAc. . .).
FT CARBOHYD 1040 1040 N-LINKED (GLCNAc. . .).
FT CONFLICT 680 680 H -> D (IN REF. 2.).
FT CONFLICT 838 840 RHH -> YTY (IN REF. 2.).
SQ SEQUENCE 1142 AA: 128352 MW: 93830687035F66 CIRC64:

```

```

Query Match 9.3%: Score 61.5; DB 1: Length 1142:
Best Local Similarity 19.7%: Pred. No. 1.2e-02:
Matches 14: Conservative 9: Mismatches 29: Indels 19: Gaps 1:

```

```

Oy 28 QXSLAISEHLKLTLCQ-----CDADVPSPGDSRLPAVDENKAGE 68
Db 610 KSTWTFKSDSPKMOQVDEHPHNTPDPKETITFMEQDPLDQDEHPFOGSRNDME 669
Oy 69 PVHLSAPAKH 79
Db 670 LSFKEPDYVRIH 680

```

```

RESULT 40
ID PMAI.DDPE STANDARD: PRT: 3591 AA.
AC P12353:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FLAMENTOUS HEMAGGLUTININ.
OS Humanella pertussis.
OC Bacteria: Proteobacteria: beta subdivision: Alcoligenaceae: Bordetella.
OC NCBI_TaxID:520:
OK 1)
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RA MEDLINE:90353839: PubMed:2389559:
RA GenBank:U00001: Humanella pertussis.
RA hamagglutinin: a protein processed from an unusually large precursor.
RT Mol. Microbiol. 4:787-800(1990).
RL 12)
RN SEQUENCE OF 1-3261 FROM N.A.
RA MEDLINE:98202284: PubMed:2393556:
RA GenBank:U00001: Humanella pertussis.
RT "Fluorescent hemagglutinin of Bordetella pertussis: nucleotide sequence and crucial role in adherence."
RL Proc. Natl. Acad. Sci. U.S.A. 86:2637-2641(1989).
CC -1- FUNCTION: EVIDENCE FOR A ROLE IN HOST-CELL BINDING AND INFECTION.
CC -1- SUBCELLULAR LOCATION: SURFACE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL consortium. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).
CC -----

```

```

DR EMBL: M60351: AAA22974.1: -.
DR EMBL: M60351: AAA22975.1: ALT_INIT.
DR EMBL: M60351: AAA22976.1: ALT_INIT.
RA Nucleotide hemagglutinin.
SQ SEQUENCE 3591 AA: 367420 MW: 877418030DE5318 CIRC64:

```

```

Query Match 9.3%: Score 61.5; DB 1: Length 3591:
Best Local Similarity 30.6%: Pred. No. 4.3e-02:
Matches 22: Conservative 13: Mismatches 28: Indels 9: Gaps 3:

```

```

Oy 26 EFXSLAISEHLK--TLQCDADVPSPGDSRLPA-----VDENK--ADEPVHLSA 76
Db 1740 EAMALVLAAGELTYKQITNRRAALIFAGSNRRLTAFVALLKIGRTIRGSDHIDNR 1799
Oy 77 TKHGFUTGDQ 88
Db 1800 TENFKALISSEVO 1811

```

```

Query Match 9.3%: Score 61.5; DB 1: Length 192:
Best Local Similarity 19.7%: Pred. No. 1.2e-02:
Matches 14: Conservative 9: Mismatches 29: Indels 19: Gaps 1:

```

```

Oy 28 QXSLAISEHLKLTLCQ-----CDADVPSPGDSRLPAVDENKAGE 68
Db 610 KSTWTFKSDSPKMOQVDEHPHNTPDPKETITFMEQDPLDQDEHPFOGSRNDME 669
Oy 69 PVHLSAPAKH 79
Db 670 LSFKEPDYVRIH 680

```

```

RESULT 41
ID VIF_HYIA2 STANDARD: PRT: 192 AA.
AC P03402:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE HIV-1 NON INFECTIVITY FACTOR (SOR PROTEIN).
OS Human immunodeficiency virus type 1 (HIV-2/SF2 isolate) (HIV-1).
OC Viruses: Retroviruses: Retroviridae: Lentivirus.
OK NCBI_TaxID:11685:
RN 1)
RN SEQUENCE FROM N.A.
RA MEDLINE:8508453: PubMed:2739227:
RA GenBank:U00001: HIV-1.
RA Stemplein W.K., Brown-Shine S.L., Gee W.W., Renard A., Randolph A., Levy J.A., Dina D., Luciv P.A., Gao W.W., Renard A., Randolph A., Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
RT Science 227:484-492(1985).
RL CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL consortium. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).
RT EMBL: K02007: AAB59877.1: -.
RT PIR: A04003: ASLJ01.
DR HIV: K02007: VIFSF2.
DR InterPro: IPR000475: Vif_1 Infect.
DR PIR: PF00559: Vif_1.
DR PRIMS: PF00349: VIRIONINFECT.
RA Protein: P000053: Viral_infect. 1.
SQ SEQUENCE 192 AA: 22460 MW: 059393039P1145B CIRC64:

```

```

Query Match 9.2%: Score 61: DB 1: Length 192:
Best Local Similarity 19.0%: Pred. No. 1:
Matches 19: Conservative 18: Mismatches 39: Indels 24: Gaps 3:

```

```

Oy 8 LILKVV-----TKVCTEALAYETQSLASMSHNLTKLDCADLVPRPSPSLAY 61
Db 8 MIVQVQRYARKITKMSLVKKNHYLSKKAKSGFYRNHYESTFVPSSEVNIPLCDALVIT 67
Oy 62 QEWGADQRYVNDSPAKIKPQILTLTDQSPKRGKSSLSSTQW 101
Db 68 TWVG-----LITGEK--WNLGSGVALIEM 89

```


COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
1-2-3, Hirosaki, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp/
URL: http://genome-gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagao, S., Sasaki
N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Genome-wide identification of transcriptionally active genes by
trehalose and its application to the isolation of full-length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2): 520-524 (1998)
Itoh, M., Katsunari, T., Shibata, Y., Shibata, K., Iwata, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (3): 463-470 (1999)
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.etc.riken.go.jp/>) for
further details.

FEATURES

SOURCE

Location/Qualifiers
1..229
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="5430596A08"
/clone_lib="RIKEN full-length enriched, adult male
oligofactory bulb"
/sex="male"
/tissue="type=olfactory brain"
/db_xref="db_xref"
/note="Site:1: Salt: Site:2: BamHI: cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
generated with primers: cDNA library was prepared with
primers: cDNA was prepared with the primer: adapter of
sequence 15' GAGGAGATTCCTCGATATTAATTAATTCGCCGCCGCC
3' cDNA was ligated into the XbaI and BamHI sites.
Vector: pGEM-3.5. The cDNA was ligated into the XbaI
from Lambda Pflc I. Cloning sites, 5' end: Salt: 3' end:
BamHI"

BASE COUNT 51 a 36 c 46 g 96 t
ORIGIN

alignment_scores:
Quality: 58.00 Length: 11
Ratio: 5.273 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 81.818

alignment_block:
US-09-471-276-831_COPY_1_16 x BB048712 ..
Align seg 1/1 to: BB048712 from: 1 to: 229

6 ValPheLeuLeuLeuTPGilyaValThrTPGily 16
|||||
178 GTTTTGTGTTGTTGGGAGTACTTGCGGC 210

seq_name: gb_gst2.BG177264

seq_documentation_block:

LOCUS

BG177264 898 bp mRNA EST 06-FEB-2001
DEFINITION 602314452F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:442061 5',
mRNA sequence.

ACCESSION

VERSION BG177264.1 GI:12683967

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

ATP80S

NIH-MGC

http://mgi.mc.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Journal.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cagbbs@femail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

NIH-MGC

http://mgi.mc.nih.gov/

DNA Sequencing: Life Technologies, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/MLM at:

http://image.jnl.gov

Plate: ULM10157 Row: 9 Column: 14

High quality sequence spot: 703.

FEATURES

SOURCE

Location/Qualifiers

1..898

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAG:442061"

/clone_lib="NIH_MGC_85"

/tissue="type=lymphoma, cell line"

/db_xref="DB108 (phage-resistant)"

/note="Organ: Lymph. Vector: PCMV-SORTG; Site:1: NCI;

Site:2: NCI; Average insert size 1.657 kb library

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC library."

BASE COUNT 189 a 277 c 289 g 143 t

ORIGIN

alignment_scores:
Quality: 56.00 Length: 12
Ratio: 5.091 Gaps: 0
Percent Similarity: 91.667 Percent Identity: 75.000

alignment_block:
US-09-471-276-831_COPY_1_16 x BG177264/rev ..
Align seg 1/1 to reverse of: BG177264 from: 1 to: 898

5 ValValPheLeuLeuLeuTPGilyaValThrTPGily 16
|||||
805 GTTGTGTTGCTGCTGTTGCGGTGCGGTGCGGCGGC 770

seq_name: gb_gst2.BG662957

seq_documentation_block:

LOCUS

BG662957 581 bp mRNA EST 30-APR-2001
DEFINITION DRA0503 Rat DRG Library Rattus norvegicus cDNA clone DRA0503 5',
mRNA sequence.

ACCESSION

VERSION BG662957.1 GI:13884879

KEYWORDS

EST.

SOURCE

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

Basal 1 to 581)

seq_name: gb_gst2.BG177264

strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized AV canal at 15 dpc library cDNA library preparation: mblee@tigr.org

1. 255 bp "Rattus sp." mblee@tigr.org
 /db.xref:"taxon:10116"
 >POLY A:Simple-report

Seq primer: M13 Forward

FEATURES
 source

Location/Qualifiers
 1..255 bp "Rattus norvegicus"
 /strain:"Sprague-Dawley"
 /db.xref:"taxon:10116"
 /clone:"UI-R-AFI-aax-d-07-0-UI"
 /clone_1db:"UI-R-AFI"
 /dex_stage:"adult"
 /lab_host:"DH10B (Life Technologies)"
 /note:"Vector: pRT33-Pac (Pharmacia) with a modified RSV promoter. The pRT33-Pac vector was constructed from the RSV library is a normalized library constructed from a 15 dpc rat atrioventricular (AV) canal. The tag is a string of 5 nucleotides present between the NotI site and the oligo-dT track. The library was constructed as described by Bonaldi, Lemmon and Soares, Genome Research 6: 791-806 (1996). Tissue provided by Jim Lin, Department of Biology, University of Iowa.
 TAC: TGGTCTTCTTGGCGAGTCTTGGCGG 18
 TAC-TISSUE-AV canal at 15 dpc
 TAC-SEB-CACGC"

BASE COUNT 48 a 30 c 10 g 48 t

ORIGIN

seq_documentation_block: 255 bp mRNA EST 17-DEC-1999

LOCUS AW253137

DEFINITION UI-R-B20-aet-a-03-0-UI.s1 UI-R-B20 Rattus norvegicus cDNA clone

ACCESSION AW253137

VERSION AW253137.1 GI:6596728

KEYWORDS EST

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus: 1 (bases 1 to 255)

REFERENCE 1 (bases 1 to 255)

APPROX 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

TITLE 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

JOURNAL 90 TTTTCTTTCTTGGCGAGTCTTGGCGG 61

COMMENT

seq_name: gb_cst1.A1010058

seq_documentation_block: 252 bp mRNA EST 15-JUN-1998

LOCUS A1010058

DEFINITION EST204509 Normalized rat lung, Bento Soares Rattus sp. cDNA clone

ACCESSION A1010058

VERSION A1010058.1 GI:3223890

KEYWORDS EST

SOURCE Rattus sp.

FEATURES
 source

Location/Qualifiers
 1..255 bp "Rattus sp." mblee@tigr.org
 /db.xref:"taxon:10116"
 /clone:"RUBPT10"
 /clone_1db:"Normalized rat lung, Bento Soares"
 /note:"Organ: lung; Vector: pRT33-Pac; Site:1: EcoRI; Site:2: NotI"

BASE COUNT 89 a 51 c 58 g 54 t

ORIGIN

seq_documentation_block: 255 bp mRNA EST 17-DEC-1999

LOCUS AW253137

DEFINITION UI-R-B20-aet-a-03-0-UI.s1 UI-R-B20 Rattus norvegicus cDNA clone

ACCESSION AW253137

VERSION AW253137.1 GI:6596728

KEYWORDS EST

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus: 1 (bases 1 to 255)

REFERENCE 1 (bases 1 to 255)

APPROX 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

TITLE 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

JOURNAL 90 TTTTCTTTCTTGGCGAGTCTTGGCGG 18

COMMENT

seq_name: gb_cst1.AW253137

seq_documentation_block: 255 bp mRNA EST 17-DEC-1999

LOCUS AW253137

DEFINITION UI-R-B20-aet-a-03-0-UI.s1 UI-R-B20 Rattus norvegicus cDNA clone

ACCESSION AW253137

VERSION AW253137.1 GI:6596728

KEYWORDS EST

SOURCE Norway rat.

FEATURES
 source

Location/Qualifiers
 1..255 bp "Rattus sp." mblee@tigr.org
 /db.xref:"taxon:10116"
 /clone:"RUBPT10"
 /clone_1db:"Normalized rat lung, Bento Soares"
 /note:"Organ: lung; Vector: pRT33-Pac; Site:1: EcoRI; Site:2: NotI"

BASE COUNT 89 a 51 c 58 g 54 t

ORIGIN

seq_documentation_block: 255 bp mRNA EST 17-DEC-1999

LOCUS AW253137

DEFINITION UI-R-B20-aet-a-03-0-UI.s1 UI-R-B20 Rattus norvegicus cDNA clone

ACCESSION AW253137

VERSION AW253137.1 GI:6596728

KEYWORDS EST

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus: 1 (bases 1 to 255)

REFERENCE 1 (bases 1 to 255)

APPROX 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

TITLE 7 PhelanLeuLeuTRPp1ValThrTPG1y 16

JOURNAL 90 TTTTCTTTCTTGGCGAGTCTTGGCGG 18

COMMENT

seq_name: gb_cst1.AW253137

seq_documentation_block: 255 bp mRNA EST 17-DEC-1999

LOCUS AW253137

DEFINITION UI-R-B20-aet-a-03-0-UI.s1 UI-R-B20 Rattus norvegicus cDNA clone

ACCESSION AW253137

VERSION AW253137.1 GI:6596728

KEYWORDS EST

SOURCE Norway rat.

```

/clone="U1-R-BU0-set-a-03-0-U1"
/clone_lib="U1-R-BU0"
/abv_stage="adult"
/abv_pos="OH08 (Life Technologies)"
/abv_pos_desc="OH08 (Life Technologies) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The U1-R-BU0
library is a subtracted library derived from the U1-R-AI,
U1-R-AB1, U1-R-AC1, U1-R-AD1, U1-R-AE1, U1-R-AF1, and
U1-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
cDNA library derived from the subtracted library and the
and the oligo-dr track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG_LIB=U1-R-BU0
TAG_RISSUE=ventricle at 16.5 dpc
TAG_SPO=GTTCG
BASE COUNT      77 a      49 c      46 g      83 t
ORIGIN
alignment_scores:
    Quality: 54.00      Length: 10
    Ratio: 5.400      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 80.000
alignment_block:
    us-09-471-276-831_COPY_1_16 x AM253137/rev ..
    Align seq 1/1 to reverse of: AM253137 from: 1 to: 255
7 Phleuleuleutrpolyvalthrtppgly 16
|||||:|||||:|||||:|||||:|||||:
90 TTTCTGTTCTTTCGCGACTTACTTGGGCG 61
seq_name: gb_cst1.A1176065
seq_documentation_block:
LOCUS      A1176065      320 bp      mRNA      EST
DEFINITION EST218642 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
AUTHORS    ROYBORG J. end. mRNA sequence.
ACCESSION  A1176065
VERSION     A1176065.1
KEYWORDS   EST
SOURCE     Rattus sp.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
           1 (bases 1 to 320)
           1 (bases 1 to 320)
REFERENCE   1 (bases 1 to 320)
AUTHORS    Chandler, I., Mason, T.M., Quackenbush, J.,
           Kerlavage, A.P. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
Unpublished (1998)
Other ESTs: TC52105
COMMENT     The Institute for Genomic Research
           4712 Research Center Drive, Rockville, MD 20850, USA
           Tel: (301)-838-3329
           Fax: (301)-838-0208
           Email: nhlee@igf.org
           Seq primer: M13-21.
           Location/Qualifiers
               1..320     =Rattus sp.
               /db_xref="taxon:10116"
               /db_xref="taxon:10118"
               /clone="U1-R-BU0"
               /clone_lib="U1-R-BU0"
               /note="Vector: p7730-Pac (Pharmacia) with a modified
               /polylinker. Site 1: Not I; Site 2: Eco RI; The U1-R-BU0
               /library is a subtracted library derived from the U1-R-AI

```

```

           /clone="U1-R-BU0"
           /clone_lib="U1-R-BU0"
           /note="Vector: p7730-Pac (Pharmacia) with a modified
           /polylinker. Site 1: Not I; Site 2: Eco RI; The U1-R-BU0
           /library is a subtracted library derived from the U1-R-AI
BASE COUNT      99 a      61 c      71 g      89 t
ORIGIN
alignment_scores:
    Quality: 54.00      Length: 10
    Ratio: 5.400      Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 80.000
alignment_block:
    us-09-471-276-831_COPY_1_16 x A1176065/rev ..
    Align seq 1/1 to reverse of: A1176065 from: 1 to: 320
7 Phleuleuleutrpolyvalthrtppgly 16
|||||:|||||:|||||:|||||:|||||:
86 TTTCTGTTCTTTCGCGACTTACTTGGGCG 57
seq_name: gb_cst1.AM956508
seq_documentation_block:
LOCUS      AM956508      364 bp      mRNA      EST
DEFINITION U1-R-E1-Fk-g-09-0-U1.s1 U1-R-E1 Rattus norvegicus cDNA clone
AUTHORS    U1-R-E1-Fk-g-09-0-U1.3' similar to g112826863emb/L02013301HS244F1
           Human DNA sequence from PAC 244F1 on chromosome 6q16.1-16.3.
           Containus nucleophosmin (nucleolar phosphoprotein, B23, NP61,
           nucleolin) like pseudogene, ESTs and an STS with a CAG repeat
           expansion polymorphism, mRNA sequence.
ACCESSION  AM956508
VERSION     AM956508.1
KEYWORDS   EST
SOURCE     Norway rat.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus norvegicus.
           1 (bases 1 to 364)
REFERENCE   1 (bases 1 to 364)
AUTHORS    Bonaldo, M.P., Lennon, G. and Soares, M.B.
           Normalization and subtraction: two approaches to facilitate gene
           discovery
           Genome Res. 6 (3), 791-806 (1996)
           97044477
           On May 7, 1998, this sequence version replaced g1:3120203.
           Program for Rat Gene Discovery and Mapping
           University of Iowa
           451 Eckstein Medical Research Building Iowa City, IA 52242, USA
           Tel: 319 335 8250
           Fax: 319 335 9565
           Email: moore@bluewebg.uoiowa.edu
           The sequence tag present in the cDNA between the NotI site and the
           3' poly(A) tail is identical to the sequence of the cDNA clone
           12-D-Subtract Library (identical to the cDNA clone 12-D-Subtract
           Bonaldo, Ph.D. Clone distribution: clones will be available through
           Research Genetics. This clone is also available through the
           I.M.A.G.E. Consortium at LML (info@image.llnl.gov). IMAGE
           ID=1719665 The following repetitive elements were found in this
           cDNA sequence: 1-32, >POLY-A1Simple-repeat
           Seq primer: M13 Forward
           POLY-A: no.
           Location/Qualifiers
               1..364     =Rattus norvegicus
               /strain="Sprague-Dawley"
               /db_xref="taxon:10116"
               /clone="U1-R-E1-Fk-g-09-0-U1"
               /clone_lib="U1-R-E1"
               /note="Vector: p7730-Pac (Pharmacia) with a modified
               /polylinker. Site 1: Not I; Site 2: Eco RI; The U1-R-BU0
               /library is a subtracted library derived from the U1-R-AI

```


The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: mtleele1@igf.org

FEATURES
Seq primer location/Qualifiers
1. 375

source
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RCover35"
/gene="Ovarian Ovary Vector: p7713pac; Site_1: EcoRI;
Site_2: NotI; 61 c 95 g 96 t

BASE COUNT 123 a 61 c 95 g 96 t

alignment_scores:
Quality: 54.00 Length: 10
Ratio: 5.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:
US-09-471-276-831_COPY_1_16 x A1408881/rev ..

Align seg 1/1 to reverse of: A1408881 from: 1 to: 375

7 PhaeLeuLeuEntpClyValThrTrpGly 16
|||||.....
74 TTTTCTGTTCTTGCGAGTTACTTGCGCG 45

seq_name: gb_est1:A1101493

seq_documentation_block:

LOCUS A1101493 379 bp mRNA EST 31-JAN-1999
DEFINITION Normalized rat brain, Bento Soares Rattus sp. cDNA clone
ACCESSION A1101493
VERSION A1101493.1 GI:3706384
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus; Rodentia; Scuriognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 379)
Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Karlavagse,A.R. and Adams M.D.
Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat

AUTHORS
Title Gene Index
Journal Input Listed (1998)
COMMENT Other Listed TC52105

JOURNAL
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: mtleele1@igf.org
Seq primer location/Qualifiers
1. 379

FEATURES
source
/organism="Rattus sp."
/db_xref="ATCC (inhost):2024108"
/db_xref="taxon:10118"
/clone="R88B87"
/gene="Ovarian Ovary Vector: p7713pac; Site_1: EcoRI;
Site_2: NotI; 61 c 93 g 103 t

BASE COUNT 122 a 61 c 93 g 103 t
ORIGIN

alignment_scores:
Quality: 54.00 Length: 10
Ratio: 5.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:
US-09-471-276-831_COPY_1_16 x A1101493/rev ..

Align seg 1/1 to reverse of: A1101493 from: 1 to: 379

7 PhaeLeuLeuEntpClyValThrTrpGly 16
|||||.....
74 TTTTCTGTTCTTGCGAGTTACTTGCGCG 45

seq_name: gb_est1:A1704632

seq_documentation_block:

LOCUS A1704632 388 bp mRNA EST 03-JUN-1999
DEFINITION U1-R-AB1-yw-d-06-0-01.s1 U1-R-AB1 Rattus norvegicus cDNA clone
U1-R-AB1-yw-d-06-0-01.3, mRNA sequence.
ACCESSION A1704632
VERSION A1704632.1 GI:4992532
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 388)
Lundquist,M.P., G. and Soares M.R.

AUTHORS
Title Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (5), 791-806 (1996)
MEDLINE 97044477
COMMENT

CONTACT: Soares, MB
Program for Rat Gene Discovery and Mapping
Department of Molecular Biology and Biophysics
451 Research Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msquares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the poly A
tail and the oligo-dT tail was used to identify clones from the
normalized ventricle at 16.5 dpc library cDNA library preparation:
M.B. Soares Lab cDNA distribution: clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-41,
->Poly.A[Simple-repeat
Seq primer: M13 Forward
POLY-A-test: M13 Forward

FEATURES
source
location/Qualifiers
1. 388

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="U1-R-AB1-yw-d-06-0-01"
/clone="106-R-AB1"
/clone="106-R-AB1"
/lab_host="m10108 (Life Technologies)"
/note="Vector: p77130-pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The U1-R-AB1
library is a normalized library constructed from 16.5 dpc
rat ventricle. The tag is a string of 5 nucleotides
present between the Not I site and the oligo-dT track.
The U1-R-AB1 library was constructed from the
Lennon and Soares' Ovarian Ovary Vector: p7713pac.
Tissue provided by Jim Liu, Department of Biology,
University of Iowa.
TAG-lib=U1-R-AB1
TAG-tissue=ventricle at 16.5 dpc

6: 791-806, 1996.
TAG-L18-UI-R-BU0
TAG-L18-UI-R-BU0
TAG-SPD-CAGC: canal at 15 dpc
156 a 76 c 104 g 134 t
ORIGIN

alignment_scores:

Quality: 54.00 Length: 10
Ratio: 5.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:

US-09-471-276-831_COPY_1_16 x AM251345/rev ..

Align seg 1/1 to reverse of: AM251345 from: 1 to: 470

7 PhdauLeuLeuTgPcUvNthTtTgCly 16
|||||
90 TTTTGTCTTGGCGAGTACTTGCGCG 61

seq_name: qb-est1.AM252673

seq_documentation_block:

LOCUS AM252673 479 bp mRNA EST 17-DEC-1999
DEFINITION UT-R-BU0-neb-C-05-0-UI-31 UT-R-BU0 Rattus norvegicus cDNA clone
ACCESSION AM252673
VERSION AM252673.1 GI:6596264
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus; Rattus; Rattus.

REFERENCE
1 (bases 1 to 475)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
AUTHORS GenBank
MEDLINE 204447
COMMENT Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9565
Email: maor@eshbuih.wesg.uiowa.edu

The sequence contained an oligo-dT track that was present in the
oligo-dT track. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized ventricle at 16.5 dpc library cDNA library preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com) The following repetitive
sequence was found in this cDNA sequence: 1-41,
>NOX.A51mple.repeat
Seq primer: M13 Forward
POLYA-Yes.

FEATURES

SOURCE Location/Qualifiers

1..475
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="ui-r-bu0-neb-c-05-0-UI-3"
/clone.lib="ui-r-bu0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT730-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-BU0
library is a subtracted library derived from the UI-R-AAL,

UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat at 16.5 dpc. Ventricle at 16.5 dpc. RV canal
at 15 dpc. Ventricle at 15 dpc. The tag is a
string of 5-6 nucleotides present between the NotI site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.

TAG-L18-UI-R-BU0
TAG-SPD-CAGC: ventricle at 16.5 dpc
159 a 76 c 106 g 134 t
ORIGIN

alignment_scores:

Quality: 54.00 Length: 10
Ratio: 5.400 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:

US-09-471-276-831_COPY_1_16 x AM252673/rev ..

Align seg 1/1 to reverse of: AM252673 from: 1 to: 475

7 PhdauLeuLeuTgPcUvNthTtTgCly 16
|||||
90 TTTTGTCTTGGCGAGTACTTGCGCG 61

seq_name: qb-est1.A1176739

seq_documentation_block:

LOCUS A1176739 479 bp mRNA EST 20-JAN-1999
DEFINITION Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
ACCESSION A1176739
VERSION A1176739.1 GI:3727377
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus; Rattus; Rattus.

REFERENCE
1 (bases 1 to 479)
Lee,N.H., Glodetz,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Ratlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
AUTHORS
TITLES
COMMENT The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel.: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21

FEATURES

SOURCE Location/Qualifiers

1..479
/organism="Rattus sp."
/db_xref="ATCC (IHOSL):2031503"
/db_xref="taxon:10118"
/clone="A1176739"
/clone.lib="Normalized rat ovary, Bento Soares"
/note="Organ: ovary; Vector: pT730Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT 159 a 77 c 107 g 136 t
ORIGIN

alignment_scores:

Quality: 54.00 Length: 10
Ratio: 5.400 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 80.000

alignment_block:

us-09-471-276-831_copy_1.16 x A1176739/rev

Align seg 1/1 to reverse of: A1176739 from: 1 to: 479

7 Phelanleuaurtpolyvalbthrtgcy 16
|||||
86 TTTTGTCTGTTGGGAGTACTGCGGG 57

seq_name: gb-ss: A2776969

seq_documentation_block: 550 bp DNA GSS 16-FEB-2001

LOCUS A2776969

DEFINITION 280011K05F Mouse 10kb plasmid U0CC1M library Mus musculus genomic

clone U0CC2M001K05 F. DNA sequence.

ACCESSION A2776969

VERSION A2776969.1 GI:12905099

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 550) Kammala, Eutharia, Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Becerra, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmood, M., Meenen, E., Petersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R., 2001. Mouse genome scaffolding with paired end reads from 10kb

plasmid inserts

COMMENT Unpublished (2000)

CONTACT: Robert B. Weiss

UNIVERSITY OF UTAH

Box 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

81201-585 5606

TEL: 801 585 7177

FAX: 801 585 7177

EMAIL: dunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0011 row: K column: 05

Seq primer: CGTGTAAACGACGCCACGT

Class: plasmid ends

High qual location/Qualifiers

1. 550

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="U0CC2M001K05"

/clone="11b-Mouse 10kb plasmid U0CC1M library"

/lab_host="F. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD207. Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnars/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was ligated to the blunt ends in high molar excess. The

polynucleotide kinase Adapter oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (9114731149b)AT129072.1), a copy number

unclonable derivative of plasmid R1. The vector was ligated

BASE COUNT 143 a 108 c 139 g 159 t 1 others

alignment_scores:

Quality: 54.00 Length: 12

Score: 54.00

Percent Similarity: 83.333 Percent Identity: 75.000

alignment_block:

us-09-471-276-831_copy_1.16 x A2776969

Align seg 1/1 to: A2776969 from: 1 to: 550

5 ValValPhelanleuaurtpolyvalbthrtgcy 16
|||||
159 GTTGTCTGCTGTGATTTGGGCGGCGACACGCGC 154

seq_name: gb-ss12: B1298800

seq_documentation_block: 433 bp mRNA EST 20-JUL-2001

DEFINITION UT-8-CV2-chn-g-10-0-UT-3', mRNA sequence.

ACCESSION UT-8-CV2-chn-g-10-0-UT-3', mRNA sequence.

VERSION B1298800

KEYWORDS B1298800.1 GI:14975080

EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 433) Kammala, Eutharia, Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS Bonaldo, M.F., Lennon, G., and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

COMMENT Program for Rat Gene Discovery and Mapping

UNIVERSITY OF IOWA

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

TEL: 319 335 8250

FAX: 319 335 9565

EMAIL: mscores@blue.wesg.uiowa.edu

UT-8-CV2-chn-g-10-0-UT-3' cDNA track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag served to verify it as a clone from the

normalized rat eye library cDNA library preparation: M.B. Soares

Lab Clone distribution: clones will be available through Research

Genetics (www.regencon.com). The following repetitive elements were

Seq primer: M13 forward

Location/Qualifiers

1. 433

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="11b-UT-R-CV2"

/clone="11b-UT-R-CV2"

/seq_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/note="Vector: p1770-Pac (Pharmacia) with a modified

polylinker. Site.1: Not I; Site.2: Eco RI; CV2 is a

unclonable derivative of plasmid R1. The vector was ligated

to the blunt ends in high molar excess. The adapted DNA was

purified and size-selected for a 9.5 to 10.5 kb range using

preparative agarose gel electrophoresis. Vector DNA was prepared

from a derivative of PMD42 (9114731149b)AT129072.1), a copy number

unclonable derivative of plasmid R1. The vector was ligated to the

[illegible]

REFERENCE	Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi: Teleostei: Cyprinodontiformes: Poeciliidae: Gambusia holbrooki			
AUTHORS	Mammals: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus			
TITLE	1 (bases 1 to 152)			
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery			
MEDLINE	Genome Res. 6 (9), 791-806 (1996)			
COMMENT	Contact: Source: MB Project: Eukaryotic Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: msotares@iuii.weeg.uiowa.edu The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA. Library Preparation: M.B. Soares lab clone distribution clones will be available through Research Genetics (www.resgen.com) Seq primer: M13 Forward POLY-A-Tes.			
FEATURES	Location/Qualifiers			
Source	1..152			
	/db_xref="taxon:9606"			
	/db_xref="taxon:10116"			
	/clone="UI-R-B51-gz-d-04-0-UI"			
	/clone="lib-UI-R-B51"			
	/db_xref="Gene"=embryonic 13 dpc			
	/db_xref="Gene"=adult with a modified			
	/db_xref="Gene"=P171D-Pac (Pharmacia)			
	Library is derived from 13 dpc whole embryo tissue. For detailed description of the library from which this clone was derived, please visit our web site at http://www.resgen.com			
	TAG_SEQ=None found"			
BASE COUNT	31 a	23 c	49 g	49 t
ORIGIN				
alignment_scores:				
Quality:	52.00	Length: 10		
Ratio:	5.778	Gaps: 0		
Percent Similarity:	90.000	Percent Identity: 80.000		
alignment_block:				
US-03-471-276-931-COVY_L16 x BR108672	..			
Align seq 1/1 to: BR108672	from: 1 to: 152			
7 PhleuleuleutufpcilyvalThrrpdy 16				
30 TTCATTCCTGCTGGGCGGCGGCGGGCA 59				
seq_name: gb_ces1.AV22375				
seq_documentation_block:				
LOCUS	AV22375 232 bp mRNA EST 05-NOV-1999			
DEFINITION	AV22375 RIKEN full-length enriched, adult male testis (DH108) Mus musculus cDNA clone A931424020 3' similar to AB023219 Homo sapiens mRNA for KIAA1002 protein, mRNA sequence.			
ACCESSION	AV22375			
KEYWORDS	EST			
KEYWORDS	EST			
KEYWORDS	house mouse			
KEYWORDS	Mus musculus			
KEYWORDS	Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi: Teleostei: Cyprinodontiformes: Poeciliidae: Gambusia holbrooki			
KEYWORDS	Mammals: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus			
KEYWORDS	1 (bases 1 to 232)			
KEYWORDS	Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Carinelli,P., Endo,T.,			
REFERENCE				
AUTHORS				

TITLE
RIKEN MOUSE ESTING PROJECT

COMMENT
CONTACT Yoshitake Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-6246
E-mail: gsc@riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Sasaki,N., Izawa,M., Matsuki,M., Ozawa,K., Tenaka,T., Yoneda,Y.,
Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

FEATURES
source
further details

location/Qualifiers

1..332
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="49314240Z"
/cds="1" ..Riken full-length enriched, adult male testis
/map="DH08"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project. The genomic clones were deposited at the National Geneclome Science Center and Genesense Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGACATAGAATCATCAACCTCTTTTTTTTAA 3']. cDNA was transcribed by using triethanolse thermo-activated reverse transcriptase and subsequently enriched for full-length by double-strand specific exonuclease digestion. The primary template cDNA was prepared with the GAGACATAGAATTCAGATTAAATTAATTAATGCCCGCCGGCCC 3'/. cDNA was cloned into the XhoI and Blnk sites. Vector: a modified plasmidscrip KS(+) after bulk excision from LambdaFLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT
ORIGIN

alignment_scores:
quality: 52.00 length: 11
ratio: 5.200 gaps: 0
Percent Similarity: 90.909 Percent Identity: 72.727

alignment_block:

alignment_scores: Quality: 52.00 Length: 16
 Ratio: 4.000 Gaps: 0
 Percent Similarity: 81.250 Percent Identity: 43.750

alignment_block:
 US-09-471-276-831_copy_1_16 x AVJ78462 ..

Align seg 1/1 to: AVJ78462 from: 1 to: 300

1 MotserleluvalaPhleuleuLeuTRPGlyValTrIPrGly 16
 145 CTCCTCTCCAGATCGTGGATCGTCGTCGGGCTCATCGGAGC 192

seq_name: gb_gsl2:B1425622

seq_documentation_block:
 LOCUS B1425622 361 bp mRNA EST 16-JUN-2001
 DEFINITION B1425622.1 cont1.1 glytine max cdna clone GENOME SYSTEMS CLONE
 ID: B1425622.1 GI:15202845
 ACCESSION B1425622
 VERSION B1425622.1 GI:15202845
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max [Fabaceae]; Strophophylla; Umbellifera; Tracheophyta; Spermatophyta; Magnoliopsida; Eudicotyledons; Core eudicotyledons; Rosales; Eurosida 1; Fabales; Papilionaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 361)
 Shoemaker, R., Keim, P., Vodkin, L., Eppelring, J., Corryall, V., Khanna, A., Boller, B., Morita, M., Hillier, L., Kneuba, T., Martin, J., Beck, C., Wylie, E., Brown, B., Schmitt, R., Siegel, R., Reising, B., Allen, M., Bowers, J., Johnson, B., Bhat, R., Johnson, R., Johnson, Y., Cardenas, M., McMan, R., Ralston, R., Kohn, S., Shiu, T., Jackson, Y., Cardenas, M., McMan, R., Ralston, R., and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 National Plant Science School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle, St. Louis, Missouri 63134 For further information call: (800) 430-0000 or (314) 427-3222 FAX: (888) 919-3224 or (314) 427-3222
 info@genomesystems.com web site: www.genomesystems.com
 Trace considered overall poor quality
 High quality sequence stop: 1.
 Location/Qualifiers
 1..361
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone_lib="GEN0375"
 /clone_lib_c="c10457"
 /library="whole seedlings of greenhouse grown plants"
 /dev_stage="3 week old"
 /db_host="DH103"

alignment_scores: Quality: 52.00 Length: 14
 Ratio: 4.333 Gaps: 0
 Percent Similarity: 85.714 Percent Identity: 57.143

alignment_block:
 US-09-471-276-831_copy_1_16 x B1425622/rev ..

Align seg 1/1 to reverse of: B1425622 from: 1 to: 361

3 MetleluvalaPhleuleuLeuTRPGlyValTrIPrGly 16
 44 TTGATTTCCTGTTCTTCCTCATGAGAGGTGTTGGGCT 3

seq_name: gb_gsl2:B1053983

seq_documentation_block:
 LOCUS B1053983 423 bp mRNA EST 15-JUN-2001
 DEFINITION B1053983.27.0101-002-g06 GN0375 Homo sapiens CDNA, mRNA sequence.
 ACCESSION B1053983
 VERSION B1053983.1 GI:14463533
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eultheria; Primates; Catarrhini; Hominoidea; Homo.
 Dias, Wido B., Garcia, Correa, R., Veloski, Almeida, S., Brites, M. R., Nagel, M. A., da Silva, W. Jr., Zazo, M. A., Bordin, S., Costa, F. P., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bal, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, P., Brentani, R. R., Reis, L. F., de Souza, J. J. and Simpson, A. J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 cDNA libraries
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 Contact: Simpson A.J.C.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil 15-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?LI-PM362-PM3-GN0375-270101-002-g06&C3-01-276&4-1)
 Seq primer: puc 18 forward
 High quality sequence stop: 421.
 High quality sequence stop: 421.
 Location/Qualifiers
 1..423
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0375"
 /dev_stage="Adult normal;
 /Site: 2; SnaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

BASE COUNT      84 a      97 c      128 g      114 t
ORIGIN
Alignment_scores:
  Quality:      52.00      Length:      14
  Ratio:        4.727      Gaps:        0
  Percent Similarity: 78.571  Percent Identity: 57.143
Alignment_block:
US-09-471-276-831_copy_1_16 x B1035983
..
Align seq 1/1 to: B1035983 from: 1 to: 423
3 MetleuValValPheIleuIleuTrpGlyValAlaThrTrpCly 16
||||| ||| :|||||:|||||:|||||:|||||
210 ATCTGTGAGTGCTGGATGATGCTGCTGGAGATCTGACGTGGCA 251

seq_name: gb_gss:A0975452
seq_documentation_block:
LOCUS      A0975452              450 bp      DNA
DEFINITION RPI-23-331110.ty RPI-23 Mus musculus genomic clone RPI-23-331110
ACCESSION  A0975452
VERSION    A0975452.1 GI:6805909
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus [Taxid:10090]
REFERENCE  1 (bases 1 to 450)
AUTHORS    Mammalla, Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            Zhao, S.; Nierman, W.; Feldblyum, T.; Malek, J.; Shatsman, S.; Alinret,
            B.; Levins, M.; McGinn, S.; Tseyaye, G.; Geer, K.; Krol, M.; de Jong, P.
            and Fraser, C.H.
            Mouse BAC end sequences from library RPI-23
            JOURNAL
            MEDLINE
            COMMENT
            Title: Mouse BAC end sequences from library RPI-23
            Journal:
            Medline:
            Comment:
            Contact: Shuying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: shuying@igmc.org
            Clones are derived from the mouse BAC library RPI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@ejg.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/cdb/bac-ends/mouse/bac_end_intro.html
            Seq primer: 77
            Class: BAC ends.
FEATURES
    source
        1..450
            /organism="Mus musculus"
            /strain="C57BL/6J 0090"
            /clone="RPI-23-331110"
            /clone_11b="RPI-23"
            /sex="Female"
            /lab_host="DH10B"
            /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1:
            EcoRI; Site: 2: EcoRI; Female C57BL/6J mouse kidney and/or
            brain genomic DNA was isolated and partially digested
            with EcoRI and EcoRV. The resulting fragments were ligated
            with selected DNA was cloned into the pBAC3.6 vector at the
            EcoRI sites. The ligation products were transformed into
            DH10B electrocompetent cells (BRL Life Technologies).
            87 a      102 c      127 g      132 t
            2 others
BASE COUNT
ORIGIN

```

```

Alignment_scores:
  Quality:      52.00      Length:      14
  Ratio:        4.000      Gaps:        0
  Percent Similarity: 92.857  Percent Identity: 57.143
Alignment_block:
US-09-471-276-831_copy_1_16 x A0975452
..
Align seq 1/1 to: A0975452 from: 1 to: 450
2 SerMetLeuValValPheIleuIleuTrpGlyValAlaThrTrp 15
:|||||:|||||:|||||:|||||:|||||
4 CTTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGACCTCG 45

seq_name: gb_gss2:B1054283
seq_documentation_block:
LOCUS      B1054283              471 bp      mRNA
DEFINITION PM3-GN0375-010201-004-b12 GN0375 Homo sapiens cDNA, mRNA sequence.
ACCESSION  B1054283
VERSION    B1054283.1 GI:14461813
KEYWORDS   EST
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 471)
            Laas, M.; E., Gutilla, Correa, R., Verjovski-Almeida, S., Brijlous, M.R.,
            Nardone, G., Gattas, M.F., Jorgensen, R., Doolittle, S., Costa, P.F.,
            Golden, G.H., Carvalho, A.F., Matsukuma, A., Baldo, G.S., Simpson, D.H.,
            Bruneau, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
            M.J., Soares, F., Brennan, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            JOURNAL
            MEDLINE
            COMMENT
            Title: Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Journal:
            Medline:
            Comment:
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704022
            Fax: +55-11-2704020
            Email: asimpson@ludwig.org.br
            This sequence was derived from the ENPES/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL:
            http://www.ludwig.org.br/scripts/gethtml2.pl?pl=PM3&lt2=PM3-GN0375-
            010201-004-b12&3=2001-02-01&4=1
            Seq primer: puc 18 forward
            High quality sequence stop: 471.
FEATURES
    source
        1..471
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_11b="GN0375"
            /over_stages="Adult"
            /note="Organ: placenta normal; Vector: puc18; Site: 1: SmaI
            products derived from ORSRES PCR (U.S. Letters Patent
            application No. 196,716 - Ludwig Institute for Cancer
            Research) profiles into the puc 18 vector. Reverse
            transcription of tissue mRNA and cDNA amplification were
            performed under low stringency conditions."
BASE COUNT      105 a      107 c      140 g      119 t
ORIGIN

```

```

alignment_block:
US-09-471-276-831_copy_1_16 x B1034203
Align seg 1/1 to: B1034203 from: 1 to: 471
3 MetlaValValPheLeuLeuTgGlyValThrTgGly 16
||||| 111 :|||||11111111111111111111
211 ATCTCTAGAGCTCGATGCTCTCGAGGACGACGAGCGGA 252
seq_name: 9b_gss:A0977668

seq_documentation_block:
LOCUS A0977668 538 bp DNA GSS 29-JAN-2000
DEFINITION RPI-23-336K8.TV RPI-23 Mus musculus genomic clone RPI-23-336K8.
ACCESSION A0977668
VERSION A0977668.1 GI:6809969
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Mus-
1 (bases 1 to 538)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akhurst,
J.B., Levine,M., McGinn,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
Mammalian BAC end sequences from library RPI-23
Unpublished (1999)
Other GSSs: RPI-23-336K8.TV
Contact: Shuying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPI-23. For BAC
library availability, please contact Pieter de Jong
(pjeterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Reson ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
Plate: 336 row: K column: 8
Seq primer: 77
Class: BAC ends.

FEATURES
Source Location/Qualifiers
1..358
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RPI-23-336K8"
/clone_id="RPI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pMAD3.6; Site:1;
EcoRI: Site.2; EcoRI: Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pMAD3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 109 a 119 c 157 g 151 t 2 others
ORIGIN
alignment_scores:
Quality: 52.00 Length: 14
Ratio: 4.000 Gaps: 0
Percent Similarity: 92.857 Percent Identity: 57.143
alignment_block:
US-09-471-276-831_copy_1_16 x A0977668

```

```

Align seg 1/1 to: A0977668 from: 1 to: 538
2 SerHeLeuValValPheLeuLeuTgGlyValThrTgGly 15
||||| 111 :|||||11111111111111111111
57 GCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 98
seq_name: 9b_gss:A2097580

seq_documentation_block:
LOCUS A2097580 704 bp DNA GSS 09-MAY-2000
DEFINITION SC10.TV RPI-23 Mus musculus genomic clone RPI-23-15C10.
ACCESSION A2097580
VERSION A2097580.1 GI:750636
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Mus-
1 (bases 1 to 704)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akhurst,
J.B., Levine,M., McGinn,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
Mouse BAC end sequences from library RPI-23
Unpublished (1999)
Other GSSs: RPI-23-15C10.TV
Contact: Shuying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPI-23. For BAC
library availability, please contact Pieter de Jong
(pjeterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Reson ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
Plate: 15 row: C column: 10
Seq primer: SP6
Class: BAC ends.

FEATURES
Source Location/Qualifiers
1..704
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RPI-23-15C10"
/clone_id="RPI-23"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pMAD3.6; Site:1;
EcoRI: Site.2; EcoRI: Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pMAD3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 133 a 164 c 204 g 203 t
ORIGIN
alignment_scores:
Quality: 52.00 Length: 14
Ratio: 4.000 Gaps: 0
Percent Similarity: 92.857 Percent Identity: 57.143
alignment_block:
US-09-471-276-831_copy_1_16 x A2097580
Align seg 1/1 to: A2097580 from: 1 to: 704

```


29 67 10.1

9	67	10.1	533	2	T38644	conserved hypothe
---	----	------	-----	---	--------	-------------------

30	66.5	10.1	44.3	2	T31441		probable glutamyl
31	66	10.0	19.2	2	S42997		viral infectivity
32	66	10.0	19.2	2	S42975		viral infectivity
33	66	10.0	19.2	2	S42958		viral infectivity
34	66	10.0	19.2	2	S42972		viral infectivity
35	66	10.0	19.2	2	S42972		viral infectivity
36	65.5	9.9	25.6	2	S34245		hypothetical prote
37	65	9.8	19.2	2	T094442		vif protein - hma
38	65	9.8	19.2	2	S42998		viral infectivity
39	65	9.8	19.2	2	S42959		viral infectivity
40	65	9.8	25.0	2	T08071		L-aspartate peroxi
41	65	9.8	28.6	2	T94369		hypothetical prote
42	65	9.8	30.7	2	G70764		probable blaC - my
43	64.5	9.8	45.3	2	B35916		site-specific reco
44	64.5	9.8	45.3	2	B35916		site-specific reco
45	64.5	9.8	25.1	2	S29747		gpi protein bac
46	64.5	9.8	25.1	2	S29747		gpi protein bac
47	64.5	9.8	74.2	2	B83834		hypothetical prote
48	64	9.7	19.2	2	S42992		viral infectivity
49	64	9.7	19.2	2	S42977		viral infectivity
50	64	9.7	19.2	2	S42970		viral infectivity
51	64	9.7	19.2	2	S42992		viral infectivity
52	64	9.7	19.2	2	T01686		viral infectivity
53	64	9.7	19.2	2	T01686		viral infectivity
54	64	9.7	33.3	2	J00256		hypothetical prote
55	64	9.7	10.0	2	T04462		hypothetical prote
56	64	9.7	10.6	2	D85383		hypothetical prote
57	63.5	9.6	25.8	2	B61228		class I histocomp
58	63.5	9.6	36.2	2	JH0288		class I histocomp
59	63.5	9.6	60.4	2	T08330		hypothetical prote
60	63.5	9.6	16.9	1	C0H04B		collagen alpha 1(I
61	63	9.5	19.2	2	S42991		viral infectivity
62	63	9.5	19.2	2	S42996		viral infectivity
63	63	9.5	19.2	2	S42996		viral infectivity
64	63	9.5	19.2	2	S42996		viral infectivity
65	63	9.5	19.2	2	S42990		viral infectivity
66	63	9.5	19.2	2	S42959		viral infectivity
67	63	9.5	19.2	2	S42959		viral infectivity
68	63	9.5	33.0	2	T46256		brevecin - human
69	63	9.5	36.6	2	B18122		lymphocyte antigen
70	63	9.5	36.6	2	I13715		MHC class I histoc
71	63	9.5	75.2	2	A65039		DNA topoisomerase
72	63	9.5	75.2	2	A65039		DNA topoisomerase
73	63	9.5	10.2	2	G85351		cellulose 1,4-beta
74	62.5	9.5	34.7	2	G72564		cellulose 1,4-beta
75	62.5	9.5	35.6	2	JH0289		class I histocomp
76	62.5	9.5	50.8	1	JCS713		5-hydroxytryptam
77	62.5	9.5	60.4	2	T08222		hypothetical prote
78	62.5	9.5	70.2	2	T31261		hypothetical prote
79	62.5	9.5	77.9	2	T14845		antifreeze-like pr
80	62.5	9.5	11.8	2	S48601		gene d protein
81	62.5	9.5	16.9	1	C0MS0B		collagen alpha 1(I
82	62	9.4	19.2	2	S43000		viral infectivity
83	62	9.4	19.2	2	S42953		viral infectivity
84	62	9.4	19.2	2	S42953		viral infectivity
85	62	9.4	31.6	2	T10435		probable transpos
86	62	9.4	36.2	2	T61997		MHC class I histoc
87	62	9.4	36.6	2	T61997		MHC class I histoc
88	62	9.4	62.0	2	JH0871		95k goli antigen
89	62	9.4	94.6	2	S71168		Ca2+-transporting
90	62	9.4	10.2	2	D86402		protein envelope C
91	62	9.4	10.2	2	D86402		protein envelope C
92							

C:Species: Mus musculus (house mouse)
 C:Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 05-Nov-1999
 C:Accession: J5895
 R:Rasmussen, J.: Fukuda, D.; Tsuji, A.; Nagabukuro, A.; Matsuda, Y.; Nishikawa, Y.; Ohya
 A:Title: Genomic structures and chromosomal location of p91, a novel murine regulatory t
 A:Reference number: J5894; MIM:38218758
 A:Accession: J5895
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Cross-references: GB:AF041035; NID:92791631; PIDN:AB95927.1; PID:92791632
 C:Comments: This protein function as inhibitory cell surface molecule against cell active
 A:Map position: 7
 F:1-23/Domain: signal sequence #status predicted <SIC>
 F:24-118/Domain: transmembrane #status predicted <TM>
 F:635-674/Domain: transmembrane #status predicted <TM>
 F:675-680/Domain: cytoplasmic #status predicted <CYT>

Query Match 11.3%; Score 74.5; DB 2; Length 680;
 Best Local Similarity 28.6%; Pred. No. 4.9;
 Matches 32; Conservative 11; Mismatches 48; Indels 21; Gaps 4;

OY 6 VELLACVTCMPVTEA---AIFETQXSLMAESHXKLTGGCCADAVCPGDSRLPAAV 61
 DB 215 VELLACVTCMPVTEA---AIFETQXSLMAESHXKLTGGCCADAVCPGDSRLPAAV 61
 OY 62 OMKAGPEPHNDSPAIRHOFLLTGDDO---GVCRCRSLGSGXQSLKLEI 110
 DB 261 KSKKSTSTOTLOOPGKGRFPISVTOOHAGQYRCVCYSAGMSGSPSLTEL 312

RESULT 6

S6692
 C:Accession: AF033
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S46492
 R:Aimes, R.T.; French, D.L.; Ouligley, J.P.
 A:Title: Cloning of a 72 kDa matrix metalloproteinase (gelatinase) from chicken embryo
 A:Reference number: S46492; MIM:34280397
 A:Accession: S46492; MIM:34280397
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-663 <AM>
 A:Cross-references: EMBL:007775; NID:9504475; PIDN:AA19596.1; PID:9504476
 A:Notes: In the authors' translation 205-Asp is shown after residue 201 and, consequently
 C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopoietin repeat homol
 C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopoietin repeat homol
 C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopoietin repeat homol
 F:230-221/Domain: fibronectin type II repeat homology <ZFP>
 F:288-329/Domain: fibronectin type II repeat homology <ZFP>
 F:346-387/Domain: fibronectin type II repeat homology <ZFP>
 F:466-663/Domain: hemopoietin repeat homology <HPN>
 F:999/400/404/410/Binding site: zinc, catalytic (His, His, His, His) (inhibited) #status
 F:401/Active site: zinc, catalytic (His) (active) #status predicted
 F:401/Active site: Gln #status predicted

Query Match 11.0%; Score 73; DB 1; Length 663;
 Best Local Similarity 23.6%; Pred. No. 6.9;
 Matches 26; Conservative 13; Mismatches 37; Indels 34; Gaps 4;

OY 17 PYTEALIFETQXSLMAESHXKLTGGCCADAVCPGDSRLPAAV 69
 DB 530 PODEKAVVF-----AGNEVWYVTSLSGKQSLKLEITGTPVYLAQS 119
 OY 70 VHLDSPAIRHOFLLTGDDOGRRCRSLGSGXQSLKLEITGTPVYLAQS 119
 DB 581 -----KRTYIFSGDWKYN-----EKKKMETLTPKFRIND 612

RESULT 7
 S4294
 C:Species: human immunodeficiency virus type 1
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999
 C:Accession: S4294
 R:Wieland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
 A:Title: In vivo genetic variability of the HIV-1 gene.
 A:Reference number: S42940
 A:Accession: S42940
 A:Status: preliminary
 A:Molecule type: DNA
 A:Cross-references: EMBL:230684; NID:9459605; PIDN:CA83164.1; PID:9459606
 C:Superfamily: AIDS vif protein

Query Match 11.0%; Score 72.5; DB 2; Length 192;
 Best Local Similarity 25.6%; Pred. No. 1.9;
 Matches 20; Conservative 12; Mismatches 37; Indels 9; Gaps 2;

OY 8 LLLACV-----TWGPTVEALIFETQXSLMAESHXKLTGGCCADAVCPGDSRLPAAV 61
 DB 8 MWQVDENRIFRMRSLVNRHIVSKAKGQFVRRHESHPRISSVHTPLDRAVIT 67
 OY 62 OKKK---ADEVYLDSPA 76
 DB 68 TWGLNTERGRNHLDDQA 85

RESULT 8

AB3033
 C:Accession: AF033
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: AB3033
 R:Stover, C.K.; Ertvin, X.Q.; Ertvin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: AB3033; MIM:20437337
 A:Accession: AB3033; MIM:20437337
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-489 <STO>
 A:Cross-references: GB:AE004903; GB:AE004901; NID:99551173; PIDN:AC08284.1; GSPB:GN
 A:Experimental source: strain PA01
 C:Superfamily: aldehyde dehydrogenase (AMO+); aldehyde dehydrogenase homology

Query Match 10.8%; Score 71.5; DB 2; Length 489;
 Best Local Similarity 27.4%; Pred. No. 7.1;
 Matches 32; Conservative 16; Mismatches 54; Indels 15; Gaps 4;

OY 16 GPTVEALIFETQXSLMAESHXKLTGGCCADAVCPGDSRLPAAV 74
 DB 67 GPAAAVYKAEVDFRHEIWDIIR-----ESCRKKELEWCAARATLES 119
 OY 75 -PAIRHOFLLTGDDOGRRCRSLGSGXQSLKLEITGTPVYLAQS 125
 DB 120 SPFAVHGVITVDSVPGKSSRYVSAIVGVATSPWFLLHT-GRSLAPALACNA 175

RESULT 9
 148471
 C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C:Accession: I48471
 R:Protein: J.D.B. Todd, J.A.: Rodriguez, N.R.: Ghosh, S.: Hogarth, P.M.: Micker, L.S.: Gelfi
 R:Title: 1996 #sequence_revision 18-Mar-1998 #text_change 17-Mar-1999
 A:Title: 1996 #sequence_revision 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 17-Mar-1999
 A:Reference number: I48471; MIM:33242399
 A:Accession: I48471
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-336 <RES>
 A:Cross-references: EMBL:X70980; NID:9311748; PIDN:CA50311.1; PID:9311749
 A:Keywords: Immunoglobulin receptor; Immunoglobulin homology
 F:128-180/Domain: Immunoglobulin homology <IMW>

Query Match 10.7%; Score 71; DB 2; Length 336;
 Best Local Similarity 25.2%; Pred. No. 5.2; 42; Indels 30; Gaps 3;

Matches 27; Conservative 8; Mismatches 42; Indels 30; Gaps 3;
 Oy 4 LVFLLMGVYMGVPTA-----AIFPEQXSLMSESHKLTGQCCADV 51
 Db 12 LVFLLMGVYMGVPTA-----AIFPEQXSLMSESHKLTGQCCADV 58
 Oy 52 PROSRPVAQENKAEPRVHLDSPATKHOFLTGDQGRKSGSL 98
 Db 59 -PDSS-----TQMFINGTAVOTSPISISVNSPDSEKQIGSS 100

RESULT 10

Fc gamma (19c) receptor high affinity - mouse
 A:Accession: A46480
 A:Reference name: high affinity 19c receptor
 C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 R:Osmen, N.; Kozak, C.A.; McKenzie, I.F.; Hogarth, P.M.
 J:Immunol. 148, 1570-1575, 1992
 A:Title: Structure and mapping of the gene encoding mouse high affinity Fc gamma RI and
 A:Reference number: A46480; MIM:92163399
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <CONS>
 A:Note: sequence extracted from NCBI backbone (NCBI:85205, NCBI:85208, NCBI:85212, NC
 J:Seaton, D.W.; Osmen, N.; Tale, B.; McKenzie, I.F.C.; Hogarth, P.M.
 J:Immunol. 144, 371-378, 1990
 A:Title: Molecular cloning and expression of the mouse high affinity Fc receptor for 19c
 A:Accession: M3511
 A:Residues: 1-404 <RES>
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Cross-references: GB:M3134; NID:9200752; PIDN:AAA0056.1; PID:9200753
 A:Superfamily: Fc gamma receptor 1; Immunoglobulin homology
 C:Keywords: Immunoglobulin receptor; Transmembrane protein
 F:127-139/Domain: Immunoglobulin homology <IMW>

Query Match 10.7%; Score 71; DB 2; Length 404;
 Best Local Similarity 25.5%; Pred. No. 6.4; 47; Indels 36; Gaps 5;

Matches 28; Conservative 9; Mismatches 37; Indels 36; Gaps 5;
 Oy 4 LVFLLMGVYMGVPTA-----AIFPEQXSLMSESHKLTGQCCADV 51
 Db 12 LVFLLMGVYMGVPTA-----AIFPEQXSLMSESHKLTGQCCADV 58
 Oy 52 PROSRPVAQENKAEPRVHLDSPATKHOFLTGDQGRKSGSL 98
 Db 59 -PDSS-----TQMFINGTAVOTSPISISVNSPDSEKQIGSS 100

RESULT 11

JC5894
 killer cell inhibitory receptor p91a precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 17-Mar-1999
 R:Watanabe, T.; Kikuchi, D.; Tsuji, A.; Nagabukuro, A.; Matsuda, Y.; O
 R:Biochem. 123, 358-368, 1998
 A:Title: Genomic structures and chromosomal location of p91, a novel murine regulator
 A:Reference number: JC5894; MIM:9821758
 A:Accession: JC5894
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-344 <RAW>
 A:Cross-references: GB:AF040946
 C:Comment: This protein function as inhibitory cell-surface molecule against cell act
 C:Genetics:
 A:Map position: 7
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-94/Product: killer cell inhibitory receptor p91a #status predicted <AMT>
 F:95-188/Domain: extracellular domain #status predicted <AMT>
 F:189-344/Domain: transmembrane domain #status predicted <AMT>
 F:345-674/Domain: cytoplasmic domain #status predicted <AMT>
 F:675-765/Domain: cytoplasmic #status predicted <AMT>

Query Match 10.5%; Score 69.5; DB 2; Length 841;
 Best Local Similarity 27.2%; Pred. No. 22; 45; Indels 25; Gaps 5;

Matches 31; Conservative 13; Mismatches 45; Indels 25; Gaps 5;
 Oy 6 VFLLMGVYMGVPTA-----AIFPEQXSLMSESHKLTGQCCADV -PCPRDSRLP 59
 Db 215 VELLVSGNLOKPTKAEKPSVITSKAMTWC-----GMLDAVYVYINSGSKTO 266
 Oy 60 AVQENKAEPRVHLDSPATKHOFLTGDQ-----GRYKNSGLSTGKSLKLEL 110
 Db 267 STGT-----LQPSKMKATPTSKRNQMKQYCTGSGKNSQPSDTLEL 312

RESULT 12

viral infectivity factor vif - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Accession: U64285
 R:Kriegelund, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuhn, J.E.
 A:Description: In vivo genetic variability of the HIV-1 gene.
 A:Reference number: U64285
 A:Accession: U64285
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-197 <RES>
 A:Cross-references: EMBL:Z30611; NID:9459464; PIDN:CA83088.1; PID:9459465
 C:Superfamily: AIDS vif protein

Query Match 10.4%; Score 69; DB 2; Length 192;
 Best Local Similarity 21.0%; Pred. No. 4.4; 37; Indels 24; Gaps 3;

Matches 21; Conservative 18; Mismatches 37; Indels 24; Gaps 3;
 Oy 8 LLMLNY-----TQPVYKALIFPEQXSLMSESHKLTGQCCADVPCPRDSRLP 61
 Db 8 MIVQVDNRKIRPMYSLVKHNHYSKATQKQYHNHRETPISSEVHPIQDARLVIT 67
 Oy 62 QENKAEPRVHLDSPATKHOFLTGDQGRKSGSLSTGW 101
 Db 68 TRNG-----LITQENK-----MILCNQSYISM 89

RESULT 13

viral infectivity factor vif - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

R:Wieland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
 Submitted to the EMBL Data Library, March 1994
 A:Description: In vivo genetic variability of the HIV-1 gene.
 A:Accession: S42996
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <MIE>
 A:Cross-references: EMBL:230687; NID:g459611; PIDN:CAAB3167.1; PID:g459612
 C:Superfamily: AIDS vif protein

Query Match 10.3%; Score 68; DB 2; Length 192;
 Best Local Similarity 22.0%; Pred. No. 5.7;
 Matches 22; Conservative 14; Mismatches 40; Indels 24; Gaps 3;

OY 8 LILMVC-----TWCPYEAIFETQXSLMASEHKLTLCCGADVPQPCDSMLAV 61
 DB 8 MIVQVDRMRIRTMISLVKHHNHSKKAKKCFYKHHSTPRISSEVHPIGLDALVYT 67
 OY 62 QEWKQRPVHIDSPAIKHOPLTGDPOGRKRSGLSTQM 101
 DB 68 TWVG-----LITGERD--VNLGQGVSTEM 89

RESULT 18
 35390
 A:Description: Infectivity factor vif - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999
 C:Accession: S42947
 R:Wieland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
 Submitted to the EMBL Data Library, March 1994
 A:Description: In vivo genetic variability of the HIV-1 gene.
 A:Accession: S42947
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <MIE>
 A:Cross-references: EMBL:230610; NID:g459460; PIDN:CAAB3087.1; PID:g459461
 C:Superfamily: AIDS vif protein

Query Match 10.3%; Score 68; DB 2; Length 192;
 Best Local Similarity 22.0%; Pred. No. 5.7;
 Matches 22; Conservative 14; Mismatches 40; Indels 24; Gaps 3;

OY 8 LILMVC-----TWCPYEAIFETQXSLMASEHKLTLCCGADVPQPCDSMLAV 61
 DB 8 MIVQVDRMRIRTMISLVKHHNHSKKAKKCFYKHHSTPRISSEVHPIGLDALVYT 67
 OY 62 QEWKQRPVHIDSPAIKHOPLTGDPOGRKRSGLSTQM 101
 DB 68 TWVG-----LITGERD--VNLGQGVSTEM 89

RESULT 19
 35391
 A:Description: vif protein - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
 C:Accession: S33981
 R:Carlhill, F.
 Submitted to the EMBL Data Library, November 1991
 A:Reference number: S33979
 A:Accession: S33981
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-192 <CAAB>
 A:Cross-references: EMBL:211530; NID:g60192; PIDN:CAA77623.1; PID:g60195
 C:Superfamily: AIDS vif protein

Query Match 10.3%; Score 68; DB 2; Length 192;
 Best Local Similarity 22.0%; Pred. No. 5.7;
 Matches 22; Conservative 14; Mismatches 40; Indels 24; Gaps 3;
 OY 8 LILMVC-----TWCPYEAIFETQXSLMASEHKLTLCCGADVPQPCDSMLAV 61
 DB 8 MIVQVDRMRIRTMISLVKHHNHSKKAKKCFYKHHSTPRISSEVHPIGLDALVYT 67
 OY 62 QEWKQRPVHIDSPAIKHOPLTGDPOGRKRSGLSTQM 101
 DB 68 TWVG-----LITGERD--VNLGQGVSTEM 89

RESULT 20
 714785
 A:Description: hypothetical protein DKFZp586N0721.1 - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14785
 R:Guttenberg, A.; Obermaier, B.; Jones, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the EMBL Data Library, August 1999
 A:Reference number: 210184
 A:Accession: T14785
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-141 <OTR>
 A:Cross-references: EMBL:AL110265
 A:Experimental source: adult uterus; clone DKFZp586N0721
 A:Note: DKFZp586N0721.1

Query Match 10.2%; Score 67.5; DB 2; Length 141;
 Best Local Similarity 42.6%; Pred. No. 4.5;
 Matches 23; Conservative 4; Mismatches 22; Indels 5; Gaps 3;
 OY 77 IKHOPLTGDPOGRKRSGLSTQMOKSKLLE--LWSKYLKSLADQAS 126
 DB 34 LHSFVLTDGV--GRICHLVGLTFTGPTSKRVHPSPGRCFLCDLAVGSS 86

RESULT 21
 725198
 A:Description: 3-oxoacid acyl synthetase component cfr4 [imported] - Pseudomonas syringae
 C:Species: Pseudomonas syringae
 C:Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jun-2000
 C:Accession: J05748
 R:Penfold, C.N.; Bender, C.L.; Turner, J.G.
 Gene 183, 167-173, 1996
 A:Reference number: J05748
 A:Accession: J05748; M010:97149295
 A:Molecule type: DNA
 A:Residues: 1-177 <PEN>
 A:Cross-references: GB:U56980; NID:g1655810; PID:g1655814
 C:Genetics: cfr4
 C:Superfamily: Pseudomonas syringae coronafaciens acid synthetase protein 4

Query Match 10.2%; Score 67.5; DB 2; Length 177;
 Best Local Similarity 39.6%; Pred. No. 5.8;
 Matches 19; Conservative 3; Mismatches 13; Indels 13; Gaps 3;
 OY 52 PQGDSRLPA--VQEWKQRPVHIDSPAIKHOPLTGDPOG--HYVC 93
 DB 125 PPDQVLPALLADMD-----LDSGASHYLVSTHDTGTHKVC 165

RESULT 22
 726320
 A:Description: hypothetical protein SCE9.01 - Streptomyces coelicolor (fragment)
 C:Species: Streptomyces coelicolor

OY 8 LLLKGV-----TMCPTVEALFYETQXSLMAESBHLKTLGQCDADVPGRPDSRLAV 61
 DB 8 MVMQVDRMRIRTMKSLVKNHMYTSKAKCMQYKIHRETHPRISSEVHPILODALVTR 67
 OY 62 QEMKQEPVHLDSPAIKIQPFLITGDPGRKRCRGSLSTQ 101
 DB 68 TWYGLQ-----TQEND--WHLQCGVSTEM 89

 RESULT 27
 542949
 viral infectivity factor vif - human immunodeficiency virus type 1
 C:Accession: U01201
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999
 R:Wieland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
 submitted to the EMBL Data Library, March 1994
 A:Description: In vivo genetic variability of the HIV-1 gene.
 A:Reference number: 542940
 A:Accession: S42949
 A:Gene: vif
 A:Molecule type: DNA
 A:Residues: 1-192 <X1>
 A:Cross-references: EMBL:230609; NID:9459462; PIDN:CAAB3086.1; PID:9459463
 C:Superfamily: AIDS vif protein

Query Match 10.1% Score 67; DB 2; Length 192;
 Best Local Similarity 21.0%; Pred. No. 73; Mismatches 40; Indels 24; Gaps 3;
 Matches 21; Conservative 15; Mismatches 40; Indels 24; Gaps 3;

 OY 8 LLLKGV-----TMCPTVEALFYETQXSLMAESBHLKTLGQCDADVPGRPDSRLAV 61
 DB 8 MVMQVDRMRIRTMKSLVKNHMYTSKAKCMQYKIHRETHPRISSEVHPILODALVTR 67
 OY 62 QEMKQEPVHLDSPAIKIQPFLITGDPGRKRCRGSLSTQ 101
 DB 68 TWYGLQ-----TQEND--WHLQCGVSTEM 89

 RESULT 28
 156130
 HLA-B*401 - human
 C:Species: Homo sapiens (man)
 C:Accession: U01201
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 R:Hildebrand, W.H.; Magfaglia, J.A.; Little, A.W.; Parham, P.
 J. Immunol. 148, 1155-1162, 1992
 A:Title: HLA-B*22: a family of molecules with identity to HLA-B7 in the alpha 1-helix.
 A:Reference number: 156130
 A:Accession: M156130; MIMD:92148136
 A:Status: preliminary
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-362 <RSS>
 A:Cross-references: GB:M77774; NID:9184116; PIDN:AA03686.1; PID:9184117
 C:Superfamily: class I histocompatibility antigen: immunoglobulin homology
 F:220-285/Domain: Immunoglobulin homology <IMM>

Query Match 10.1% Score 67; DB 2; Length 362;
 Best Local Similarity 24.2%; Pred. No. 15; Mismatches 64; Indels 22; Gaps 5;
 Matches 32; Conservative 14; Mismatches 64; Indels 22; Gaps 5;

 OY 6 VPLLKGV-----TMCPTVEALFYETQXSLMAESBHLKTLG-----QCDADVPGR 52
 DB 9 LHLMLGALALATETVKNSSHMITYT-TAMSRGKCHNRYANGVYDQVQVFTSPASFP 67
 OY 53 PCDSRLAVQESKQEPVHLDSPAIKIQPFLITGDPGRKRCRGSLSTQWQSLLELTG 112
 DB 68 RGEPAFAPVDEGS--PEYWDHNTQYKAAQOATDRESLEMLNR--GYNOSDAQSHWT 119
 OY 113 PVLKAGSLADG 124

DB 120 QTRVGGDCLDPCG 131

 RESULT 29
 conserved hypothetical protein SPAC323.07c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: J38644
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Boche, G.; Kampsperger, U.; Pohl, T.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: J38644
 A:Accession: J38644
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-533 <NO>
 A:Cross-references: EMBL:AL109988; PIDN:CA85410.1; GSPDB:GN00066; SPDB:SPAC323.07c
 A:Experimental source: strain 972h; cosmid c323
 C:Genetics:
 A:Gene: SPDB:SPAC323.07c
 A:Molecule type: DNA
 A:Map position: 1
 A:Interon: 221/2

Query Match 10.1% Score 67; DB 2; Length 533;
 Best Local Similarity 22.9%; Pred. No. 24; Mismatches 38; Conservative 16; Mismatches 62; Indels 50; Gaps 7;
 Matches 38; Conservative 16; Mismatches 62; Indels 50; Gaps 7;

 OY 1 MSWLVEYLILGVYMG-----PYTFAIFETQXSLMAESBHLKTLGQCDADVPGR 53
 DB 23 LNTLNTLLVNHPTIGCFGLAPAVAAATF-----WFSQ--TCLILYICFSSFTIPMD 303

 OY 54 GDSR-----LPAVQESKQEPVHLDSPAIKIQPFLITGDT--- 87
 DB 304 GRSKQALKLSLPMILRSPHOMLKTVTBMATNYSLSAGQVLAQSGLTSLTSLP 363
 OY 88 -----GGRYRCSSGISMGCKQSLSKLRLGPKVLAIGALDA 125
 DB 364 QTPFAFVASTFVCHLIGSRANLRLGRVAVSLALCISIFDPS 409

RESULT 30
 325435
 probable glutamyl-L-PRNA reductase (EC 1.2.1.-) hema - Helicobacterium mobilis
 C:Species: Helicobacterium mobilis
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
 C:Accession: J31441
 R:Xiong, J.; Inoue, K.; Bauer, C.E.
 Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998
 A:Title: Tracking molecular evolution of photosynthesis by characterization of a major
 A:Reference number: J31441
 A:Accession: J31441; MIMD:95061195
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <X1>
 A:Cross-references: EMBL:AF080002; NID:93820536; PID:93820540; PIDN:AA04013.1
 C:Genetics:
 A:Gene: hema
 C:Superfamily: glutamyl-L-PRNA reductase
 C:Keywords: oxidoreductase

Query Match 10.1% Score 66.5; DB 2; Length 443;
 Best Local Similarity 26.2%; Pred. No. 22; Mismatches 34; Indels 47; Gaps 9;
 Matches 34; Conservative 15; Mismatches 34; Indels 47; Gaps 9;

 OY 18 VPEALFYETQXSLMAESBHLKTLG-----TGGQD--ADVAGPGRDSS--LPAVQ--- 62
 DB 17 VREKISTEADLS--EALYKLGQMGLEGCISLTGCTNTEFTIGASDIEMDKQMTAVKAFV 73
 OY 63 -ENGAQEP-----VILDSPAIKIQP-----LTGQDQ--GRYR-----C 93
 DB 74 LEMGQLOPQDSKRFYVHTLTDALIRHLRFAASGSDSVNLGDTIIGQVRYATQNSCNDIC 133

OY 94 RSGLSSTGWXQ 103
DB 134 SNGLVNTWQ 143

RESULT 31

viral infectivity factor vif - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42972 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Submitted to the EMBL Data Library, March 1994

A:Description: In vivo genetic variability of the HIV-1 gene.

A:Reference number: S42940

A:Accession: S42997

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-192 <ME>

C:Cross-references: EMBL:230688; NID:9459613; PIDN:CAAB3168.1; PID:9459614

C:Superfamily: AIDS vif protein

Query Match 10.0%; Score 66; DB 2; Length 192;
Best Local Similarity 21.0%; Pred. No. 9.3; Mismatches 21; Conservative 16; Indels 24; Gaps 3;

OY 8 LLLKGV-----TWGPTGMAIFETGXSLMSESHXKLTGCDADVGRPGDSRLPAY 61

DB 8 MIVQVDRMRIRFTWMSLVKHNHYISGKAKMCMYKHNHSTPRISSEVNIPLDAAVLIT 67

OY 62 QEWGAEVPHNDSPAIKHOFLITGDOGRKRSCLSTGW 101

DB 68 TWVG-----LHTGERD--WHLGGCVSIEW 89

RESULT 32

viral infectivity factor vif - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42972 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Submitted to the EMBL Data Library, March 1994

A:Description: In vivo genetic variability of the HIV-1 gene.

A:Reference number: S42940

A:Accession: S42975

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-192 <ME>

C:Cross-references: EMBL:230636; NID:9459514; PIDN:CAAB3113.1; PID:9459515

C:Superfamily: AIDS vif protein

Query Match 10.0%; Score 66; DB 2; Length 192;
Best Local Similarity 21.0%; Pred. No. 9.3; Mismatches 21; Conservative 15; Indels 24; Gaps 3;

OY 8 LLLKGV-----TWGPTGMAIFETGXSLMSESHXKLTGCDADVGRPGDSRLPAY 61

DB 8 MIVQVDRMRIRFTWMSLVKHNHYISGKAKMCMYKHNHSTPRISSEVNIPLDAAVLIT 67

OY 62 QEWGAEVPHNDSPAIKHOFLITGDOGRKRSCLSTGW 101

DB 68 TWVG-----LHTGERD--WHLGGCVSIEW 89

RESULT 33

viral infectivity factor vif - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42954

C:Submitted to the EMBL Data Library, March 1994

A:Description: In vivo genetic variability of the HIV-1 gene.

A:Reference number: S42940

A:Accession: S42954

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-192 <ME>

C:Cross-references: EMBL:230615; NID:9459472; PIDN:CAAB3092.1; PID:9459473

C:Superfamily: AIDS vif protein

Query Match 10.0%; Score 66; DB 2; Length 192;
Best Local Similarity 20.0%; Pred. No. 9.3; Mismatches 20; Conservative 17; Indels 24; Gaps 3;

OY 8 LLLKGV-----TWGPTGMAIFETGXSLMSESHXKLTGCDADVGRPGDSRLPAY 61

DB 8 MIVQVDRMRIRFTWMSLVKHNHYISGKAKMCMYKHNHSTPRISSEVNIPLDAAVLIT 67

OY 62 QEWGAEVPHNDSPAIKHOFLITGDOGRKRSCLSTGW 101

DB 68 TWVG-----LHTGERD--WHLGGCVSIEW 89

RESULT 34

viral infectivity factor vif - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Accession: S43006; S42948

C:Submitted to the EMBL Data Library, March 1994

A:Description: In vivo genetic variability of the HIV-1 gene.

A:Reference number: S42940

A:Accession: S43006

A:Status: preliminary

A:Molecule type: DNA

C:Cross-references: EMBL:230661; NID:9459631; PIDN:CAAB3133.1; PID:9459632; EMBL:2306

C:Superfamily: AIDS vif protein

Query Match 10.0%; Score 66; DB 2; Length 192;
Best Local Similarity 21.0%; Pred. No. 9.3; Mismatches 21; Conservative 15; Indels 24; Gaps 3;

OY 8 LLLKGV-----TWGPTGMAIFETGXSLMSESHXKLTGCDADVGRPGDSRLPAY 61

DB 8 MIVQVDRMRIRFTWMSLVKHNHYISGKAKMCMYKHNHSTPRISSEVNIPLDAAVLIT 67

OY 62 QEWGAEVPHNDSPAIKHOFLITGDOGRKRSCLSTGW 101

DB 68 TWVG-----LHTGERD--WHLGGCVSIEW 89

RESULT 35

viral infectivity factor vif - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Sep-1999

C:Accession: S42972

C:Submitted to the EMBL Data Library, March 1994

A:Description: In vivo genetic variability of the HIV-1 gene.

A:Reference number: S42940

A:Accession: S42972

A:Status: preliminary

A:Molecule type: DNA

C:Cross-references: EMBL:230633; NID:9459508; PIDN:CAAB3110.1; PID:9459509

C:Superfamily: AIDS vif protein

```

Query Match          10.0%: Score 66; DB 2; Length 192;
Best Local Similarity 20.0%: Pred. No. 9.3;
Matches 20: Conservative 17; Mismatches 39; Indels 24; Gaps 3;

OY 8 LILKCV-----TWGVTGALIFETQXSJLMESEKHLKTLGGCAVGRPRGDSRLP 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 MIVQVDRNRIRITKMSLVKHNHNTISKAKKMSYRNHSTPIRSSVNIPLDGAHLYIT 67

OY 62 OEQKADRPVNDSPAIKRIQFLITGDTGCRGRSGSLSTOW 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 TWVG-----LITGEND--WHLGCGVSIEVM 89

RESULT 36
S34245
Hypothetical protein 3 - actinophaga vmb
C>Date: 19-Mar-1995 #sequence_revision 19-Mar-1997 #extl_change 08-Oct-1999
C:Accession: S34245
C:Date: 16-Jul-1993 #sequence_revision 16-Jul-1999 #extl_change 26-Aug-1999
R:Arne, J.; van Mollert, L.; Piten, P.; Ophthakker, G.; Joris, B.; Gyssens, H.
Submitted to the EMBL Data Library, May 1993
A:Description: Nucleotide sequence analysis of the coding region containing the 3 main
A:Reference number: S34241
A:Accession: S34245
A:Molecule type: DNA
A:Residues: 1-156 <WIE>
A:Cross-references: EMBL.X20923; NID:g311847; PIDD:CA50986.1; PID:g311852
A:Superfamily: Aids vif protein

Query Match          9.9%: Score 65.5; DB 2; Length 256;
Best Local Similarity 21.0%: Pred. No. 15;
Matches 22: Conservative 17; Mismatches 47; Indels 19; Gaps 4;

OY 4 LWFFLLKLVTKNPVTR--NALPFTQXSJLME-----SEHXLTKTGCD-----A 47
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 5 LIAVYVNDQPTGPPDDVFAEYARIRGAMHMDDTGCELEIGRPFAGFADPRTDSCG 64

OY 48 DVNCPGDSRLPVAQWEG--AOEPIVNDSPAIKRIQFLITGDTGCR 89
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 DVNAPRSGDSGLDMRSTLADHETINDVADMSREKIRALAEQDC 109

RESULT 37
T09442
vif protein - human immunodeficiency virus type 1 (strain JRFL)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1993 #sequence_revision 16-Jul-1999 #extl_change 26-Aug-1999
R:Phong S.; Vinters, H.V.; Akashi, T.; O'Brien, M.A.; Chen, T.S.; Koyanagi, Y.; Nemerio,
Submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09442
A:Status: preliminary
A:Stature: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-256 <WIE>
A:Cross-references: EMBL.U63632; NID:g1465777; PID:g1465780
A:Gene: vif
C:Superfamily: Aids vif protein

Query Match          9.8%: Score 65; DB 2; Length 192;
Best Local Similarity 21.0%: Pred. No. 12;
Matches 21: Conservative 15; Mismatches 40; Indels 24; Gaps 3;

OY 8 LILKCV-----TWGVTGALIFETQXSJLMESEKHLKTLGGCAVGRPRGDSRLP 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 MIVQVDRNRIRITKMSLVKHNHNTISKAKKMSYRNHSTPIRSSVNIPLDGAHLYIT 67

OY 62 OEQKADRPVNDSPAIKRIQFLITGDTGCRGRSGSLSTOW 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 TWVG-----LITGEND--WHLGCGVSIEVM 89

```

```

DB 68 TWVG-----LITGEND--WHLGCGVSIEVM 89

RESULT 38
S42998
vif protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #extl_change 20-Sep-1999
C:Accession: S42998
R:Wietland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
Submitted to the EMBL Data Library, March 1994
A:Description: In vivo genetic variability of the HIV-1 gene.
A:Reference number: S42998
A:Accession: S42998
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <WIE>
A:Cross-references: EMBL.Z30689; NID:g459615; PIDD:CA63169.1; PID:g459616
A:Superfamily: Aids vif protein

Query Match          9.8%: Score 65; DB 2; Length 192;
Best Local Similarity 21.0%: Pred. No. 12;
Matches 21: Conservative 15; Mismatches 40; Indels 24; Gaps 3;

OY 8 LILKCV-----TWGVTGALIFETQXSJLMESEKHLKTLGGCAVGRPRGDSRLP 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 MIVQVDRNRIRITKMSLVKHNHNTISKAKKMSYRNHSTPIRSSVNIPLDGAHLYIT 67

OY 62 OEQKADRPVNDSPAIKRIQFLITGDTGCRGRSGSLSTOW 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 TWVG-----LITGEND--WHLGCGVSIEVM 89

RESULT 39
S42950
vif infectivity factor vif - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #extl_change 20-Sep-1999
C:Accession: S42950
R:Wietland, U.; Hartmann, J.; Suhr, H.; Salzberger, B.; Eggers, H.J.; Kuehn, J.E.
Submitted to the EMBL Data Library, March 1994
A:Description: In vivo genetic variability of the HIV-1 gene.
A:Reference number: S42950
A:Accession: S42950
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <WIE>
A:Cross-references: EMBL.Z30620; NID:g459482; PIDD:CA63097.1; PID:g459483
A:Superfamily: Aids vif protein

Query Match          9.8%: Score 65; DB 2; Length 192;
Best Local Similarity 21.0%: Pred. No. 12;
Matches 21: Conservative 15; Mismatches 40; Indels 24; Gaps 3;

OY 8 LILKCV-----TWGVTGALIFETQXSJLMESEKHLKTLGGCAVGRPRGDSRLP 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 MIVQVDRNRIRITKMSLVKHNHNTISKAKKMSYRNHSTPIRSSVNIPLDGAHLYIT 67

OY 62 OEQKADRPVNDSPAIKRIQFLITGDTGCRGRSGSLSTOW 101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 TWVG-----LITGEND--WHLGCGVSIEVM 89

RESULT 40
T08071
L-ascorbate peroxidase (EC 1.11.1.11) - leaf mustard
C:Species: Brassica juncea (leaf mustard)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #extl_change 19-May-2000
C:Accession: T08071
R:Toy, E.; Pua, E.C.
Submitted to the EMBL Data Library, December 1997

```

A:Reference number: 216333
 A:Accession: T08071
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-250 <FRN>
 A:Keywords: cytochrome-c peroxidase
 C:Keywords: cytochrome-c peroxidase
 F:42/Active site: His (distal axial ligand) #status predicted
 F:163/Binding site: heme iron (His) (proximal axial ligand) #status predicted
 F:179,208/Active site: Trp, Asp #status predicted

Query Match
 Best Local Similarity 29.8%; Pred. No. 15;
 Matches 28; Conservative 5; Mismatches 41; Indels 20; Gaps 5;

48 DYVPGPDSRLPAVWENGAQEPVHLDSPAIRKIQELTG-----TQGRKCR-SQLS 98
 DB 121 DKRPGPDSRLPAVWENGAQEPVHLDSPAIRKIQELTG-----TQGRKCR-SQLS 176
 0Y 99 TCKXKSLKLF-----ITGPKYLACSLALGCA 125
 DB 177 GAKTSNPLIFDINSYFKELTGKESKGLQVSDKA 210

RESULT 41
 F43936
 hypothetical protein B1D1.230 (imported) - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
 C:Accession: T43936
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
 A:Reference number: 225022
 A:Accession: T43936
 A:Molecule type: DNA
 A:Residues: 1-286 <SCCH>
 A:Cross-references: EMBL:AL359927; GSPDB:GN00116; NCSP:B1D1.230
 A:Experimental source: BAC clone B1D1, strain OR74A
 C:Genetics:
 A:Gene: NCSP:B1D1.230
 A:Map position: 6 159/1
 A:Superfamily: Neurospora crassa hypothetical protein B1D1.230

Query Match
 Best Local Similarity 9.8%; Score 65; DB 2; Length 286;
 Matches 22; Conservative 5; Mismatches 28; Indels 10; Gaps 3;

0Y 27 TQXSNSEHXKTLQCDNHPGPGSRPVP-----QNGC--ADPPHLSRAIK 78
 DB 106 TQRLVYVSAORAKADG--GVEVGTGTSRVPATVTCILSLGCAFPQTPAHISKVAVR 163
 0Y 79 HQFL 83
 DB 164 WKVL 168

RESULT 42
 G70764
 probable biac - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: G70764
 R:Conrad, R.; Davies, R.; Parthill, J.; Gantler, T.; Church, C.; Harris, D.; Gordon, S.; Concor, R.; Davies, R.; Davlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holtz, S.; Rajendram, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID: 98295987
 A:Accession: G70764
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-307 <SCCH>
 A:Keywords: beta-lactamase I
 C:Keywords: beta-lactamase I
 C:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: biac
 C:Superfamily: beta-lactamase I

Query Match
 Best Local Similarity 27.9%; Pred. No. 21;
 Matches 36; Conservative 17; Mismatches 44; Indels 32; Gaps 7;

0Y 12 GYVWGPYFEALAYE--TQXSL-----NASESHKTLQD-----OCADAPV----- 50
 DB 128 GHTIQCICDAIYSDSTANILLADLCGGCGTAFTGSLSDIVSRDLMEPELNR 187
 0Y 51 GPRGDS--LPVWENGAQEPVHLDSPAIRKIQELTG-----DTQKTRKCSGLSTCKX 103
 DB 188 DDPDEORTTPPAALAVIQVLYGNAIPPRKALITPDMAKNTTGAKRTIACGPADM-- 245
 0Y 104 LSKLEITG 112
 DB 246 --KVIDKT 252

RESULT 43
 B56916
 site-specific recombinase Int - Streptomyces lividans
 C:Species: Streptomyces lividans
 C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
 C:Accession: B56916
 R:Brasch, M.A.; Feltus, G.S.; Lee, S.C.; Cohen, S.N.
 A:Title: Localization and nucleotide sequences of genes mediating site-specific recombination
 A:Reference number: A36916; MUID: 93259953
 A:Accession: B56916
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-455 <BRAS>
 A:Note: sequence extracted from NCBI backbone (NCBIN:131954, NCBI:P:131957)

Query Match
 Best Local Similarity 30.0%; Pred. No. 32;
 Matches 21; Conservative 8; Mismatches 31; Indels 10; Gaps 2;

0Y 5 VELLNGVT-WQPTFAALPYETQXSL-----NASESHKTLQCDNAPGPG 54
 DB 338 LKVLNHNRYALVAPRDLNRTQKGLIDTGTGVNMAKSRALTPNOCASLAKRT 197
 0Y 55 DRLPAVQEM 64
 DB 398 DLKHNAAVSTM 407

RESULT 44
 K78920
 gene 20 protein - human herpesvirus 3
 C:Species: human herpesvirus 3, varicella-zoster virus
 C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: B27343
 R:Davidson, A.J.; Scott, J.E.
 J:Gen. Virol. 67, 1159-1161, 1986
 A:Title: Localization and nucleotide sequences of genes mediating site-specific recombination
 A:Reference number: A27343; MUID: 65006557
 A:Accession: B27343
 A:Molecule type: DNA
 A:Residues: 1-483 <DAV>
 A:Cross-references: EMBL:X04370; NID: g59989; PIDN: CAN27903.1; PID: g60009

C:Genetics:
A:Gene: 20
C:Superfamily: varicella-zoster virus gene 20 protein
C:Keywords: capsid assembly; DNA binding

Query Match	9.88;	Score 65;	DB 1;	Length 483;
Best Local Similarity	31.48;	Pred. No. 35;		
Matches 16;	Conservative 9;	Mismatches 26;	Indels 0;	Gaps 0;

QY 56 SKLPALVDEGAQEPPVHLDSIPAIAKHQFLLTCTDQGRYRCRSGSTGWXQLSK 106
: : : : : : : : : : : : : : : :
Db 297 ANIPACVFMDVDKDLHLSADGLKHVFLVFYYTQRQRQREGVRHLALSQLE 347

RESULT 45

ypwN protein - *Bacillus subtilis*
N: Alternate names: hypothetical protein lpa-92r

```
CISpecies: Bacillus subtilis
CIDate: 10-Sep-1999 #sequence
CRevision: 10-Sep-1999 #text_change
CAccession: E38747..E70056

```

Classification: 537.91, 570.00
 Rclasser, P.: Kunst, F.; Arnaud, M.; Coudart, M.F.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Pannont, G.; Panchin, A.

A:Title: *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from
R:Reference number: S39655; MUID:95020537
D:Accession: 330747

A:Accession: 539747
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Reg100res: 1-258 <C1A>
A:Cross-references: EMBL:X73124; NID:g413923; PIDN:CAAS1648.1; PID:g414016
A:Note: The nucleotide sequence was submitted to the NID Data Library June 1993

A.: Enright, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. *Nature* 390, 249-256, 1997

Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fume, S.; Gallizzi, A.; Gallucci, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.H.

A: Authors; Lauber, V.; Lee, S.-M.; Levine, A.; Liu, H.; Masuda, S.; Mauee, S.; Roetter, P.; Kolnigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, J.

A. Authors: Rivolta, C.; Roche, E.; Rose, M.; Saddle, Y.; Sato, T.; Scanlon
V., M.; Ugawa, N.; Ugueta, B.; Pail, S.H.; Parro, V.; Ponl, T.M.; Portetelli
Rlegger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Saddle, Y.; Sato, T.; Scanlon

T.: Winters, P.; Miyaji, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, T.; Akouchi, M.; Tamakoshi, O.; Tanaka, T.; Terpestre, P.; Togonoh, A.; Tosato, V.; Uchiyamama, S.

A: Authors: Yoshikawa, H. F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A: Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A:Reference number: A69580; MUID:98044033
A:Accession: P70056

A: status: preliminary; nucleic acid sequence not shown; translation not shown
A: molecule type: DNA
A: molecule type: RNA
A: molecule type: protein
A: molecule type: other

A: cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PID:CAB15789.1; PID:g26362968
A: cross-references source: strain 168

NCBI GenBank accession: [SRR100](#)
 C:Gene: [WTFN](#)
 C:Gene: [WTFN](#)
 C:Gene: [WTFN](#)

C:Superfamily: hypothetical protein ylb0

Query Match 9.88; Score 64.5; DB 1; Length 258;

Best Local Similarity 29.18; Pred. No. 19;
Matches 25; Conservative 14; Mismatches 20;
Indels 19; Gaps 6;

12 CATTGCPVTEAAIFYEIOXSL-MAESEHXLKRTLCQCDADVPGPFGDSRL-----PAY----- 61

Db 50 GFRNAVVRHQ--YEKALQLAKKQKQRMALGN-----GQPAKKRLLYQPPAVDPEIT 101

62 QEWGAEVHLDSPATKH-QPLLTGD 86
|||.....|....
|||.....|....

Db 102 QETAAEPPVKTETPSVENEDPLMSGE 127

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2002, 16:45:13 ; Search time 19.31 seconds
(without alignments)

146,837 Million cell updates/sec

Title: US-09-471-276-831

Perfect score: 661

Sequence: 1 KSHLWPELLMGWTCVTE.....LLEITPRYLACSLADGAS 126

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents, AA:*

1: /cgn2.6/pdata2/2/1aa/5A.COMB.pep:*

2: /cgn2.6/pdata2/2/1aa/5B.COMB.pep:*

3: /cgn2.6/pdata2/2/1aa/6A.COMB.pep:*

4: /cgn2.6/pdata2/2/1aa/6B.COMB.pep:*

5: /cgn2.6/pdata2/2/1aa/6C.COMB.pep:*

6: /cgn2.6/pdata2/2/1aa/6D.COMB.pep:*

7: /cgn2.6/pdata2/2/1aa/6E.COMB.pep:*

8: /cgn2.6/pdata2/2/1aa/6F.COMB.pep:*

9: /cgn2.6/pdata2/2/1aa/6G.COMB.pep:*

10: /cgn2.6/pdata2/2/1aa/6H.COMB.pep:*

11: /cgn2.6/pdata2/2/1aa/6I.COMB.pep:*

12: /cgn2.6/pdata2/2/1aa/6J.COMB.pep:*

13: /cgn2.6/pdata2/2/1aa/6K.COMB.pep:*

14: /cgn2.6/pdata2/2/1aa/6L.COMB.pep:*

15: /cgn2.6/pdata2/2/1aa/6M.COMB.pep:*

16: /cgn2.6/pdata2/2/1aa/6N.COMB.pep:*

17: /cgn2.6/pdata2/2/1aa/6O.COMB.pep:*

18: /cgn2.6/pdata2/2/1aa/6P.COMB.pep:*

19: /cgn2.6/pdata2/2/1aa/6Q.COMB.pep:*

20: /cgn2.6/pdata2/2/1aa/6R.COMB.pep:*

21: /cgn2.6/pdata2/2/1aa/6S.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	87.5	13.2	631	4	US-09-345-468-12
2	75.5	11.4	244	3	US-09-135-782-4
3	74.5	11.2	651	4	US-08-985-950-22
4	72.5	11.0	244	4	US-09-193-191-4
5	68.5	10.4	390	2	US-08-979-924-18
6	68.5	10.4	415	1	US-08-444-528-37
7	66.5	10.0	415	1	US-08-959-774-10
8	65.5	9.9	503	4	US-08-857-076-102
9	65	9.8	685	2	US-08-878-989-1
10	65	9.8	685	2	US-08-136-282-2
11	65	9.8	685	2	US-09-272-795-1
12	63	9.8	685	2	US-08-985-950-16
13	63	9.8	685	2	US-08-985-950-16
14	64	9.7	249	4	US-09-345-468-5
15	64	9.7	319	4	US-09-345-468-5
16	64	9.7	319	4	US-09-345-468-5
17	63.5	9.6	431	4	US-08-985-950-14
18	63.5	9.6	431	4	US-08-985-950-14
19	63.5	9.6	615	4	US-08-985-950-16
20	63.5	9.6	615	4	US-08-985-950-16
21	62.5	9.5	389	4	US-09-071-224-28
22	62.5	9.5	405	3	US-09-071-224-28
23	62.5	9.5	405	3	US-09-071-224-28
24	62.5	9.5	405	3	US-09-071-224-28
25	62.5	9.5	405	3	US-09-071-224-28
26	61.5	9.3	303	4	US-09-111-730-2
27	61.5	9.3	303	4	US-09-111-730-2
28	61.5	9.3	389	4	US-09-071-224-22
29	61.5	9.3	389	4	US-09-071-224-24
30	61.5	9.3	389	4	US-09-071-224-25
31	61.5	9.3	389	4	US-09-071-224-27
32	61.5	9.3	389	4	US-09-071-224-28
33	61.5	9.3	389	4	US-09-071-224-29
34	61.5	9.3	389	4	US-09-071-224-30
35	61.5	9.3	389	4	US-09-071-224-31
36	61.5	9.3	389	4	US-09-071-224-32
37	61.5	9.3	389	4	US-09-071-224-33
38	61.5	9.3	389	4	US-09-071-224-34
39	61.5	9.3	389	4	US-09-071-224-35
40	61.5	9.3	389	4	US-09-071-224-36
41	61.5	9.3	389	4	US-09-071-224-37
42	61.5	9.3	389	4	US-09-071-224-38
43	61.5	9.3	389	4	US-09-071-224-39
44	61.5	9.3	389	4	US-09-071-224-40
45	61.5	9.3	389	4	US-09-071-224-41
46	61.5	9.3	389	4	US-09-071-224-42
47	61.5	9.3	389	4	US-09-071-224-43
48	61.5	9.3	389	4	US-09-071-224-44
49	61.5	9.3	389	4	US-09-071-224-45
50	61.5	9.3	389	4	US-09-071-224-46
51	61.5	9.3	389	4	US-09-071-224-47
52	61.5	9.3	389	4	US-09-071-224-48
53	61.5	9.3	389	4	US-09-071-224-49
54	61.5	9.3	389	4	US-09-071-224-50
55	61.5	9.3	389	4	US-09-071-224-51
56	61.5	9.3	389	4	US-09-071-224-52
57	61.5	9.3	389	4	US-09-071-224-53
58	61.5	9.3	389	4	US-09-071-224-54
59	61.5	9.3	389	4	US-09-071-224-55
60	61.5	9.3	389	4	US-09-071-224-56
61	61.5	9.3	389	4	US-09-071-224-57
62	61.5	9.3	389	4	US-09-071-224-58
63	61.5	9.3	389	4	US-09-071-224-59
64	61.5	9.3	389	4	US-09-071-224-60
65	61.5	9.3	389	4	US-09-071-224-61
66	61.5	9.3	389	4	US-09-071-224-62
67	61.5	9.3	389	4	US-09-071-224-63
68	61.5	9.3	389	4	US-09-071-224-64
69	61.5	9.3	389	4	US-09-071-224-65
70	61.5	9.3	389	4	US-09-071-224-66
71	61.5	9.3	389	4	US-09-071-224-67
72	61.5	9.3	389	4	US-09-071-224-68
73	61.5	9.3	389	4	US-09-071-224-69
74	61.5	9.3	389	4	US-09-071-224-70
75	61.5	9.3	389	4	US-09-071-224-71
76	61.5	9.3	389	4	US-09-071-224-72
77	61.5	9.3	389	4	US-09-071-224-73
78	61.5	9.3	389	4	US-09-071-224-74
79	61.5	9.3	389	4	US-09-071-224-75
80	61.5	9.3	389	4	US-09-071-224-76
81	61.5	9.3	389	4	US-09-071-224-77
82	61.5	9.3	389	4	US-09-071-224-78
83	61.5	9.3	389	4	US-09-071-224-79
84	61.5	9.3	389	4	US-09-071-224-80
85	61.5	9.3	389	4	US-09-071-224-81
86	61.5	9.3	389	4	US-09-071-224-82
87	61.5	9.3	389	4	US-09-071-224-83
88	61.5	9.3	389	4	US-09-071-224-84
89	61.5	9.3	389	4	US-09-071-224-85
90	61.5	9.3	389	4	US-09-071-224-86
91	61.5	9.3	389	4	US-09-071-224-87
92	61.5	9.3	389	4	US-09-071-224-88
93	61.5	9.3	389	4	US-09-071-224-89
94	61.5	9.3	389	4	US-09-071-224-90
95	61.5	9.3	389	4	US-09-071-224-91
96	61.5	9.3	389	4	US-09-071-224-92
97	61.5	9.3	389	4	US-09-071-224-93
98	61.5	9.3	389	4	US-09-071-224-94
99	61.5	9.3	389	4	US-09-071-224-95
100	61.5	9.3	389	4	US-09-071-224-96


```

; APPLICANT: XU, SHUANG-YONG
; TITLE OF INVENTION: METHOD FOR CLONING THE NSPHI RESTRICTION-MODIFICATION
; TITLE OF INVENTION: SYSTEM IN E. COLI AND PRODUCING THE RECOMBINANT NSPHI
; TITLE OF INVENTION: RESTRICTION ENDONUCLEASE
; CURRENT FILING DATE: 1998-11-17
; CURRENT FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 4
; LENGTH: 244
; ORGANISM: Homo sapiens
; ORGANISM No. 6130078tcc sp.
US-09-193-191-4

```

```

Query Match 11.0%; Score 72.5; DB 4; Length 244;
Best Local Similarity 37.5%; Pred. No. 0.19; Indels 11; Gaps 4;
Matches 21: Conservative 6; Mismatches 18; Indels 11; Gaps 4;
OY 55 DSRFVAVOMAGQAEVYHDSPA---IKHQP-ILTYGTCGYYR-----CRSGSLTGC 100
DB 132 DHRFP-MERWQASEPPLTMSDDEIKOKFOLLKKAQSGNNILKMSCEKICRTG 186

```

```

RESULT 5
US-08-979-424-1
; Sequence 1, Application US/08979424
; Patent No. 5942606
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeli
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: VIRAL RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: 2000012000 Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,424
; FILING DATE: Filed Herewith.
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; FILING DATE:
; FILING DATE:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0405 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-955-0555
; TELEFAX: 650-945-4186
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; type: amino acid
; STRANDEDNESS: single
; STRANDEDNESS: single
; IMMEDIATE SOURCE:
; LIBRARY: LUNGFET03
; CLONE: 1232054
US-08-979-424-1

```

```

Query Match 10.4%; Score 68.5; DB 2; Length 390;
Best Local Similarity 24.6%; Pred. No. 1.2;
Matches 29: Conservative 15; Mismatches 43; Indels 31; Gaps 4;
OY 37 KXITLQGDADYPPGPGOSTLPAVQENGAQGVHDS---PAIKQP----- 81
DB 140 HSIKTL-ELNVLPAPPCRLGCVPHVGNANTLSQSPNSKRVAYQVQNDQQLPSRTFF 198
OY 82 -----LITGDTGFRYRGRGLSTGKXQLSKLEL-TGPKYIACSLALD 124
DB 199 APALDVIRGSLTLNLSSMAGVYCKAHNEGYACQCNVLTVEVSGCAVAVGAVG 256

```

```

RESULT 6
US-09-113-750A-38
; Sequence 38, Application US/09113750A
; Patent No. 6294176
; GENERAL INFORMATION:
; APPLICANT: David E. Junker and Mark D. Cochran
; TITLE OF INVENTION: Recombinant Raccoonpox virus
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,750A
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 55744
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)262-0400
; TELEFAX: (212)264-0525
; TELETYPE: 42523
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; type: amino acid
; STRANDEDNESS: single
; STRANDEDNESS: single
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; MOLECULE TYPE: DNA (genomic)
US-09-113-750A-38

```

```

Query Match 10.1%; Score 67; DB 4; Length 254;
Best Local Similarity 29.9%; Pred. No. 1;
Matches 26: Conservative 6; Mismatches 25; Indels 30; Gaps 5;
OY 21 MAIPEYQSIAMSRXIKTLTGCCADYNG---PBG---DSRLAVQEV----- 64
DB 63 MAIYERGSIMKRGH--DRGDDHDEGLPWPVGLSSGLSLTSYMANLAFLSYSTA 120
OY 65 -----GQEPYHLDSPAIKHOFL 82
DB 121 QHPPEITGIGWASKVYVD---IKHOPI 144

```

```

RESULT 7
US-08-464-523B-27
; Sequence 27, Application US/08464523B
; Patent No. 5723761
; GENERAL INFORMATION:

```

APPLICANT: Toni A. Voelker
APPLICANT: Ling Yuan
APPLICANT: Jean Kridl
APPLICANT: Deborah Hawkins
APPLICANT: Aubrey Jones
APPLICANT: Michael Aoyi ACP thioesterase
TITLE OF INVENTION: Sequences
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA: 08/09/464,523B
APPLICATION NUMBER: 08/251,695
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13131
FILING DATE: 10-NOV-1994
PRIOR APPLICATION DATA: 08/152,004
APPLICATION NUMBER: 08/251,695
FILING DATE: 10-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/251,695
FILING DATE: 16-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ronald E. Schmedler
ADDRESS: 4719
CITY: Santa Rosa
STATE: CA
COUNTRY: USA
ZIP: 95404
REFERENCE/DOCKET NUMBER: 36,924
TELEPHONE/DOCKET INFORMATION:
TELEPHONE: (916) 753-6313
FAX: (916) 553-5110
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS: 27:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-464-523B-27

Query Match 10.0%; Score 66; DB 1; Length 415;
Best Local Similarity 30.8%; Pred. No. 2.8;
Matches 24; Conservative 11; Mismatches 15; Indels 28; Gaps 5;

OY 69 PVLHDSPAIR-----HQP-LITSTQGRKRSQSTGSMQLSK-----LLE 109
DB 273 FLVDSVYVLESDLVKHYKRVATQDS-----LQGLTFQMDLDVNHQHSNWKTIWLE 327
OY 110 LCGPVYL---AGSLAD 123
DB 328 SMPTVLEFQSLSTATE 345

RESULT 8
US-09-999-774A-10
Sequence 10; Application US/08999774A
Patent No. 6274312
GENERAL INFORMATION:
APPLICANT: Gish, Kurt C.
APPLICANT: Scgezzi, Wolfgang
APPLICANT: Shanahan, Frances
APPLICANT: Lees, Emma K.

APPLICANT: McClanahan, Terrill K.
TITLE OF INVENTION: Intracellular Regulatory Molecules;
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,774A
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US 60/032,818
APPLICATION NUMBER: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Chung, Edwin P.
REFERENCE/DOCKET NUMBER: DX0646
TELEPHONE/DOCKET INFORMATION:
TELEPHONE: (650) 952-9358
FAX: (650) 956-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-999-774A-10

Query Match 9.9%; Score 65.5; DB 4; Length 503;
Best Local Similarity 23.7%; Pred. No. 4.2;
Matches 32; Conservative 14; Mismatches 48; Indels 41; Gaps 5;

OY 7 FLILMGVTPYVEAFYETDQSLWASRSHKATYTGQCDADVP---GP- 53
DB 105 FVLHDSQDADPHKHYALSRCLCMRSGYTDQTEKPAVAPPGCGPPPYLAHTPA 164
OY 54 -----GDSRLPVAQV-----WCADEPVHLSPAIKHIFL-LTGD 86
DB 165 GDAQGFPLPAGADKGLLDLVQVDSCLHLLTASNG--DPVFTKANDEQSDQLPQLTA 223
OY 87 TGGRYRCRSGISTCW 101
DB 224 CAGGCPCLPAGELYCW 238

RESULT 9
US-09-999-776-102
Sequence 102; Application US/08857076C
Patent No. 6225120
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
APPLICANT: Kimura, Koutarou
APPLICANT: Patterson, Garth
APPLICANT: Ogy, Scott
APPLICANT: Ogy, Scott
APPLICANT: Tlisenbaum, Heidi
APPLICANT: Morris, Jason
APPLICANT: Allison
TITLE OF INVENTION: THERAPY AND DIAGNOSTIC TOOLS FOR
TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
FILE REFERENCE: 00786/351001
CURRENT APPLICATION NUMBER: US/08/957,076C

```

US-08-878-989-1
CLONE: 39043
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 40 KTLGQ-CGAD-----VPRPGDSYELAAVQMGQEPVNDSPAIKIQP-----81
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 82 LITDGTGCGTSCSL--STCKXDLSEKLEIGTGVYVLSL 120
DB 72 LVDPTGKTKSGVGLAKGFGACSTEDTITNKVYAAKI 112
RESULT 11
US-09-136-282-2
1 Sequence 2, Application US/09136282
2 Patent No. US-6,082,082
3 OTHER INFORMATION:
4 OTHER INFORMATION:
5 APPLICANT: ANDERSON, KAREN
6 APPLICANT: JACKSON, JEFFREY
7 APPLICANT: HANSHURY, MICHAEL
8 APPLICANT: NERURKAR, SANDHYA
9 APPLICANT: KOSHIK, AMY
10 APPLICANT: RODZYK, MARK
11 TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)
12 NUMBER OF SEQUENCES: 1
13 ADDRESS: Rainer & Prestia
14 CORRESPONDENCE ADDRESS:
15 STREET: P.O. Box 980
16 CITY: Valley Forge
17 STATE: PA
18 COUNTRY: USA
19 ZIP CODE: 19082
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Diskette
22 OPERATING SYSTEM: DOS
23 SOFTWARE: FASTEST for Windows Version 2.0
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/09/136,282
26 FILING DATE: 20-AUG-1997
27 CLASSIFICATION: 20-AUG-1998
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 60/056,112
30 FILING DATE: 20-AUG-1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Prestia, Paul P
33 REGISTRATION NUMBER: 33,031
34 PENDING/NOT PENDING: GR-70231
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 610-407-0700
37 TELEFAX: 610-407-0700
38 TELEX: 846169
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 685 amino acids
42 STRANDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 US-09-136-282-2
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 40 KTLGQ-CGAD-----VPRPGDSYELAAVQMGQEPVNDSPAIKIQP-----81
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 82 LITDGTGCGTSCSL--STCKXDLSEKLEIGTGVYVLSL 120
DB 72 LVDPTGKTKSGVGLAKGFGACSTEDTITNKVYAAKI 112
RESULT 11
US-09-136-282-2
1 Sequence 2, Application US/09136282
2 Patent No. US-6,082,082
3 OTHER INFORMATION:
4 OTHER INFORMATION:
5 APPLICANT: ANDERSON, KAREN
6 APPLICANT: JACKSON, JEFFREY
7 APPLICANT: HANSHURY, MICHAEL
8 APPLICANT: NERURKAR, SANDHYA
9 APPLICANT: KOSHIK, AMY
10 APPLICANT: RODZYK, MARK
11 TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)
12 NUMBER OF SEQUENCES: 1
13 ADDRESS: Rainer & Prestia
14 CORRESPONDENCE ADDRESS:
15 STREET: P.O. Box 980
16 CITY: Valley Forge
17 STATE: PA
18 COUNTRY: USA
19 ZIP CODE: 19082
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Diskette
22 OPERATING SYSTEM: DOS
23 SOFTWARE: FASTEST for Windows Version 2.0
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/09/136,282
26 FILING DATE: 20-AUG-1997
27 CLASSIFICATION: 20-AUG-1998
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 60/056,112
30 FILING DATE: 20-AUG-1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Prestia, Paul P
33 REGISTRATION NUMBER: 33,031
34 PENDING/NOT PENDING: GR-70231
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 610-407-0700
37 TELEFAX: 610-407-0700
38 TELEX: 846169
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 685 amino acids
42 STRANDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 US-09-136-282-2
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 40 KTLGQ-CGAD-----VPRPGDSYELAAVQMGQEPVNDSPAIKIQP-----81
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 82 LITDGTGCGTSCSL--STCKXDLSEKLEIGTGVYVLSL 120
DB 72 LVDPTGKTKSGVGLAKGFGACSTEDTITNKVYAAKI 112
RESULT 11
US-09-136-282-2
1 Sequence 2, Application US/09136282
2 Patent No. US-6,082,082
3 OTHER INFORMATION:
4 OTHER INFORMATION:
5 APPLICANT: ANDERSON, KAREN
6 APPLICANT: JACKSON, JEFFREY
7 APPLICANT: HANSHURY, MICHAEL
8 APPLICANT: NERURKAR, SANDHYA
9 APPLICANT: KOSHIK, AMY
10 APPLICANT: RODZYK, MARK
11 TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)
12 NUMBER OF SEQUENCES: 1
13 ADDRESS: Rainer & Prestia
14 CORRESPONDENCE ADDRESS:
15 STREET: P.O. Box 980
16 CITY: Valley Forge
17 STATE: PA
18 COUNTRY: USA
19 ZIP CODE: 19082
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Diskette
22 OPERATING SYSTEM: DOS
23 SOFTWARE: FASTEST for Windows Version 2.0
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/09/136,282
26 FILING DATE: 20-AUG-1997
27 CLASSIFICATION: 20-AUG-1998
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 60/056,112
30 FILING DATE: 20-AUG-1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Prestia, Paul P
33 REGISTRATION NUMBER: 33,031
34 PENDING/NOT PENDING: GR-70231
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 610-407-0700
37 TELEFAX: 610-407-0700
38 TELEX: 846169
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 685 amino acids
42 STRANDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 US-09-136-282-2
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 40 KTLGQ-CGAD-----VPRPGDSYELAAVQMGQEPVNDSPAIKIQP-----81
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 82 LITDGTGCGTSCSL--STCKXDLSEKLEIGTGVYVLSL 120
DB 72 LVDPTGKTKSGVGLAKGFGACSTEDTITNKVYAAKI 112
RESULT 11
US-09-136-282-2
1 Sequence 2, Application US/09136282
2 Patent No. US-6,082,082
3 OTHER INFORMATION:
4 OTHER INFORMATION:
5 APPLICANT: ANDERSON, KAREN
6 APPLICANT: JACKSON, JEFFREY
7 APPLICANT: HANSHURY, MICHAEL
8 APPLICANT: NERURKAR, SANDHYA
9 APPLICANT: KOSHIK, AMY
10 APPLICANT: RODZYK, MARK
11 TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)
12 NUMBER OF SEQUENCES: 1
13 ADDRESS: Rainer & Prestia
14 CORRESPONDENCE ADDRESS:
15 STREET: P.O. Box 980
16 CITY: Valley Forge
17 STATE: PA
18 COUNTRY: USA
19 ZIP CODE: 19082
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Diskette
22 OPERATING SYSTEM: DOS
23 SOFTWARE: FASTEST for Windows Version 2.0
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/09/136,282
26 FILING DATE: 20-AUG-1997
27 CLASSIFICATION: 20-AUG-1998
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 60/056,112
30 FILING DATE: 20-AUG-1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Prestia, Paul P
33 REGISTRATION NUMBER: 33,031
34 PENDING/NOT PENDING: GR-70231
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 610-407-0700
37 TELEFAX: 610-407-0700
38 TELEX: 846169
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 685 amino acids
42 STRANDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 US-09-136-282-2
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 40 KTLGQ-CGAD-----VPRPGDSYELAAVQMGQEPVNDSPAIKIQP-----81
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28; Gaps 6;
Oy 82 LITDGTGCGTSCSL--STCKXDLSEKLEIGTGVYVLSL 120
DB 72 LVDPTGKTKSGVGLAKGFGACSTEDTITNKVYAAKI 112
RESULT 11
US-09-136-282-2
1 Sequence 2, Application US/09136282
2 Patent No. US-6,082,082
3 OTHER INFORMATION:
4 OTHER INFORMATION:
5 APPLICANT: ANDERSON, KAREN
6 APPLICANT: JACKSON, JEFFREY
7 APPLICANT: HANSHURY, MICHAEL
8 APPLICANT: NERURKAR, SANDHYA
9 APPLICANT: KOSHIK, AMY
10 APPLICANT: RODZYK, MARK
11 TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)
12 NUMBER OF SEQUENCES: 1
13 ADDRESS: Rainer & Prestia
14 CORRESPONDENCE ADDRESS:
15 STREET: P.O. Box 980
16 CITY: Valley Forge
17 STATE: PA
18 COUNTRY: USA
19 ZIP CODE: 19082
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Diskette
22 OPERATING SYSTEM: DOS
23 SOFTWARE: FASTEST for Windows Version 2.0
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/09/136,282
26 FILING DATE: 20-AUG-1997
27 CLASSIFICATION: 20-AUG-1998
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 60/056,112
30 FILING DATE: 20-AUG-1997
31 ATTORNEY/AGENT INFORMATION:
32 NAME: Prestia, Paul P
33 REGISTRATION NUMBER: 33,031
34 PENDING/NOT PENDING: GR-70231
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 610-407-0700
37 TELEFAX: 610-407-0700
38 TELEX: 846169
39 INFORMATION FOR SEQ ID NO: 2:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 685 amino acids
42 STRANDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: protein
45 US-09-136-282-2
Query Match
DB 20 OALGGCGAGDSSKKKRRDPREPSORPDSO--AQP-----PAAPNNINSHSCEISR 71
Matches 27: Conservative 14; Mismatches 32; Indels 28;
```



```

? TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes:
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESS: DNAX Research Institute
? 3000 Campus Drive
? Palo Alto, California 94304-1104
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/985,950
? FILING DATE: 05-DEC-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/041,279
? FILING DATE: 21-MARCH-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/033,181
? FILING DATE: 16-DEC-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/032,252
? FILING DATE: 05-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Ching, Edwin P
? REGISTRATION NUMBER: 34,090
? REFERENCE/DOCKET NUMBER: DX0670K
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650)952-9196
? TELEFAX: (650)956-1204
? INFORMATION FOR THE PUBLISHER:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 431 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-985-950-20

Query Match          9.6% Score 63.5; DB 4; Length 431;
Best Local Similarity 25.6% Pred. No. 6.2;
Matches 33; Conservative 18; Mismatches 57; Indels 21; Gaps 5;

OY 1 MSALVFLILMGVYWGPTBNAIFETQXSLNAE-----SEHIKRLTLCQCDADVGR 52
DB 18 MPTVYVLICGLSLGPRHVOGTITPFLTAMEDSVITQGSFVTLSCQSLDAD----- 73
OY 53 PGRSLPAAVOE---WGQDEYILDSALIKIQGLTGTGCRYSGLSTSGKQSLK--L 107
DB 74 --EYLYREKKSASWITRIPRLPVKNQGFHISITWEHTGYCOYASRANSELSDEPLY 131

OY 108 LEITG--PK 114
DB 132 LVMTQATPK 140

RESULT 19
US-08-985-950-16
? Sequence 16, Application US/08985950
? Patent No. 6140076
? GENERAL INFORMATION:
? TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes:
? ADDRESS: DNAX Research Institute
? 3000 Campus Drive
? Palo Alto, California 94304-1104
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30

```

```

? COUNTRY: USA
? ZIP: 94304-1104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/985,950
? FILING DATE: 05-DEC-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/041,279
? FILING DATE: 21-MARCH-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/033,181
? FILING DATE: 16-DEC-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/032,252
? FILING DATE: 05-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Ching, Edwin P
? REGISTRATION NUMBER: 34,090
? REFERENCE/DOCKET NUMBER: DX0670K
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650)952-9196
? TELEFAX: (650)956-1204
? INFORMATION FOR THE PUBLISHER:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 615 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-985-950-16

Query Match          9.6% Score 63.5; DB 4; Length 615;
Best Local Similarity 25.6% Pred. No. 10;
Matches 33; Conservative 18; Mismatches 57; Indels 21; Gaps 5;

OY 1 MSALVFLILMGVYWGPTBNAIFETQXSLNAE-----SEHIKRLTLCQCDADVGR 52
DB 18 MPTVYVLICGLSLGPRHVOGTITPFLTAMEDSVITQGSFVTLSCQSLDAD----- 73
OY 53 PGRSLPAAVOE---WGQDEYILDSALIKIQGLTGTGCRYSGLSTSGKQSLK--L 107
DB 74 --EYLYREKKSASWITRIPRLPVKNQGFHISITWEHTGYCOYASRANSELSDEPLY 131

OY 108 LEITG--PK 114
DB 132 LVMTQATPK 140

RESULT 20
US-08-985-950-18
? Sequence 18, Application US/08985950
? Patent No. 6140076
? GENERAL INFORMATION:
? TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes:
? ADDRESS: DNAX Research Institute
? 3000 Campus Drive
? Palo Alto, California 94304-1104
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30

```



```

Query Match          9.5%  Score 62.5; DB 4; Length 425;
Best Local Similarity 26.7%  Pred. No. 8.2;
Matches 32; Conservative 10; Mismatches 57; Indels 21; Gaps 5;

OY 14 TWGPT-----EAAIFETQXSLMAESKHL-----KTLGCGCADVGPFGDSRLPAY 61
DB 132 TWGHSCHTPEIDALF--TPREIWAATNLSARSDVLTLDLVYTTDPPDPAVSRY 249
OY 62 OMKAGPEPHLDSPAIRKHQELTGDQGRKRGSGTGMKQSLFLELTGKYLACSLA 121
DB 230 GGLGDLQSVRWSPALQKFLFOAKYQIRRVED--SDVMKVVDDVSNQF-----SCLLA 302

RESULT 26
US-09-111-730-2
: Sequence 2, Application US/0911730
: Patent No. 6371359
: GENERAL INFORMATION:
: APPLICANT: Hideharu Anazawa
: APPLICANT: Hitoko Shimada
: APPLICANT: Seiji Sugimoto
: APPLICANT: Shoji Saitoh
: APPLICANT: Tatsuo Sudo
: APPLICANT: Yuzuru Ishimura
: APPLICANT: Matsuhiko Hayashi
: APPLICANT: Toshiaki Monkawa
: APPLICANT: Tadashi Yoshida
: APPLICANT: Shin Maehiro
: APPLICANT: Masahiko Maehiro
: APPLICANT: Hiroyoshi Tanaka
: TITLE OF INVENTION: 25-HYDROXYVITAMIN D3-1a-HYDROXYLASE AND DNA ENCODING THE HYDROXYL
: FILE REFERENCE: 1074
: CURRENT APPLICATION NUMBER: US/09/111,730
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patent In Ver. 2.0
: STRANDNESS: single
: LENGTH: 508
: TYPE: PRT
: ORGANISM: human
US-09-111-730-2

Query Match          9.5%  Score 62.5; DB 4; Length 508;
Best Local Similarity 24.5%  Pred. No. 10;
Matches 34; Conservative 15; Mismatches 35; Indels 55; Gaps 8;

OY 13 VMGCVTEAIFETQXSLMAESKHLTTLGCGCADVGPFGDSRLPAY 60
DB 15 VMHPELGASLGYR-----EYHSARRSL---ADIPGSPSTFLAEFCCKGLSRLE 63
OY 61 VQ-----EMKAGPEPHLDSPAIRKHQEL-----LTGDTQGRKSGS- 95
DB 64 LQVQGAHFGPWLASFGVRYTVVAKALVDELHMGDEPFGPFGCSPTWTEH- RCKOR 122
OY 96 --GLSTG---KXOLSKLL 108
DB 123 AGCLITAEKESKORLSRL 141

RESULT 27
US-09-071-224-23
: Sequence 23, Application US/09071224
: Patent No. 6271343
: GENERAL INFORMATION:
: APPLICANT: Lok, Si
: APPLICANT: Presnell, Scott R.
: APPLICANT: Jelmeberg, Anna C.
: APPLICANT: Gilbert, Teresa
: APPLICANT: Foster, Donald C.
: APPLICANT: Adams, Robyn L.
: APPLICANT: Lehner, Joyce M.
: TITLE OF INVENTION: MAMMALIAN ZCYTORS

```

```

: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Zymogenetics
: STREET: 1201 Eastlake Ave East
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/071,224
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: REGISTRATION NUMBER: 32,743
: REFERENCE/DOCKET NUMBER: 96-22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-442-6657
: TELEFAX: 206-442-6678
: TEXT:
: INFORMATION FOR SEQ ID NO: 23:
: DEFINITION: ZCYTORS
: LENGTH: 303 amino acids
: TYPE: amino acid
: STRANDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-071-224-23

Query Match          9.3%  Score 61.5; DB 4; Length 303;
Best Local Similarity 26.7%  Pred. No. 6.9;
Matches 32; Conservative 10; Mismatches 57; Indels 21; Gaps 5;

OY 14 TWGPT-----EAAIFETQXSLMAESKHL-----KTLGCGCADVGPFGDSRLPAY 61
DB 152 TWGHSCHTPEIDALF--TPREIWAATNLSARSDVLTLDLVYTTDPPDPAVSRY 209
OY 62 OMKAGPEPHLDSPAIRKHQELTGDQGRKRGSGTGMKQSLFLELTGKYLACSLA 121
DB 210 GGLGDLQSVRWSPALQKFLFOAKYQIRRVED--SDVMKVVDDVSNQF-----SCLLA 262

RESULT 28
US-09-071-224-22
: Sequence 22, Application US/09071224
: Patent No. 6271343
: GENERAL INFORMATION:
: APPLICANT: Lok, Si
: APPLICANT: Presnell, Scott R.
: APPLICANT: Jelmeberg, Anna C.
: APPLICANT: Gilbert, Teresa
: APPLICANT: Foster, Donald C.
: APPLICANT: Adams, Robyn L.
: APPLICANT: Lehner, Joyce M.
: TITLE OF INVENTION: MAMMALIAN ZCYTORS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Zymogenetics
: STREET: 1201 Eastlake Ave East
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:

```

```

? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: DOS
? CURRENT APPLICATION DATA: Windows Version 2.0
? APPLICATION NUMBER: US/09/071.224
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? ATTORNEY/AGENT INFORMATION:
? NAME: Lunn, Paul G
? REGISTRATION NUMBER: 32.743
? REFERENCE/DOCKET NUMBER: 96-22
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 206-442-6627
? TELEFAX: 206-442-6678
?
? INFORMATION FOR SEQ ID NO: 22:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 389 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
? US-09-071-224-22

```

Query Match 9.3%; Score 61.5; DB 4; Length 389;
 Best Local Similarity 26.7%; Pred. No. 9.8;
 Matches 32; Conservative 10; Mismatches 57; Indels 21; Gaps 5;

```

OY 14 TACGPT-----EAAIFETQXSIWAASEHKL-----KTLGGCCADAVRPPGDSRLPAY 61
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 132 TWCHPSCHIPKDALP--TVEIWEATNRILNSANSVDVLTLDLVYTTDPDPDVNSRV 209
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 62 QEMKAGEPVHDSPAIKHOFELTGDTQGRKRSGLSTGMXSLKLELTGPKVYLACSLA 121
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 210 GLELDQLSVRWSPALPKDFLEOKAYQIRFVED--SDVMKVVDDVSNQT-----SCHLA 262
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 29
 US-09-071-224-24
 Sequence 24, Application us/09071224
 Patent No. 6271343
 GENERAL INFORMATION:
 APPLICANT: LOK, SI
 APPLICANT: PRESNELL, SCOTT H.
 APPLICANT: JELBERG, ANNA C.
 APPLICANT: GILBERT, TERESA
 APPLICANT: FOSTER, DONALD C.
 APPLICANT: ADAMS, ROBYN L.
 APPLICANT: LEHNER, JOYCE M.
 TITLE OF INVENTION: MAMMALIAN ZCYTORS
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ZymoGenetics
 STREET: 1201 Eastlake Ave East
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/071.224
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:

```

? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Lunn, Paul G
? REGISTRATION NUMBER: 32.743
? REFERENCE/DOCKET NUMBER: 96-22
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 206-442-6627
? TELEFAX: 206-442-6678
?
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 389 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
? US-09-071-224-24

```

Query Match 9.3%; Score 61.5; DB 4; Length 389;
 Best Local Similarity 26.7%; Pred. No. 9.8;
 Matches 32; Conservative 10; Mismatches 57; Indels 21; Gaps 5;

```

OY 14 TACGPT-----EAAIFETQXSIWAASEHKL-----KTLGGCCADAVRPPGDSRLPAY 61
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 132 TWCHPSCHIPKDALP--TVEIWEATNRILNSANSVDVLTLDLVYTTDPDPDVNSRV 209
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 62 QEMKAGEPVHDSPAIKHOFELTGDTQGRKRSGLSTGMXSLKLELTGPKVYLACSLA 121
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 210 GLELDQLSVRWSPALPKDFLEOKAYQIRFVED--SDVMKVVDDVSNQT-----SCHLA 262
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 30
 US-09-071-224-25
 Sequence 25, Application us/09071224
 Patent No. 6271343
 GENERAL INFORMATION:
 APPLICANT: LOK, SI
 APPLICANT: PRESNELL, SCOTT R.
 APPLICANT: JELBERG, ANNA C.
 APPLICANT: GILBERT, TERESA C.
 APPLICANT: FOSTER, DONALD C.
 APPLICANT: ADAMS, ROBYN L.
 APPLICANT: LEHNER, JOYCE M.
 TITLE OF INVENTION: MAMMALIAN ZCYTORS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ZymoGenetics
 STREET: 1201 Eastlake Ave East
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/071.224
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Paul G
 REGISTRATION NUMBER: 32.743
 REFERENCE/DOCKET NUMBER: 96-22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-442-6627
 TELEFAX: 206-442-6678

OY 62 QEWKQEPVHDSRAIKHIDELTGTGGRKRSQSGTGMXQSLLELTGPKVLACSLA 121
DB 213 GGLHDSLSYRWVSPALPKDFLPQAKYQIRRVED--SYDMKVVDDVSNQ-----SCLRLA 265

RESULT 33
US-09-071-224-4

? Sequence 4, Application US/09071224
? Patent No. 6271343
? GENERAL INFORMATION:
? APPLICANT: Presnell, Scott R.
? APPLICANT: Jeleberry, Anna C.
? APPLICANT: Gilbert, Teresa
? APPLICANT: Foster, Donald C.
? APPLICANT: Adams, Robyn L.
? APPLICANT: Lehner, Joyce W.
? APPLICANT: HANSON, KAMALIAN ZCYTON5
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Zymogenetics
? STREET: 1201 Eastlake Ave East
? CITY: Seattle
? STATE: WA
? COUNTRY: USA
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSD for Windows Version 2.0
? CURRENT APPLICATION DATA:
? FILING DATE: US/09/071.224
? PRIORITY NUMBER:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: LUMI, PAUL G.
? REGISTRATION NUMBER: 32.743
? REFERENCE/DOCKET NUMBER: 96-22
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 206-442-6627
? TELEFAX: 206-442-6678
? TRILER:
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 425
? TYPE: amino acids
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: Protein
? FRAGMENT TYPE: Internal
? US-09-071-224-4

Query Match 9.3%, Score 61.5, DB 4, Length 425;
Best Local Similarity 26.7%, Pred No. 11;
Matches 32, Conservative 10, Mismatches 57, Indels 21, Gaps 5;
OY 14 TWQPT-----EAAIFETQSLMASSHKL-----KTGGCAGDVGVGPPDSRLAV 61
DB 188 TWQPSHSHPKDLALF--TVEIWEATNLSARSADVLTDLIDVYTTDPDPVSHV 245
OY 62 QEWKQEPVHDSRAIKHIDELTGTGGRKRSQSGTGMXQSLLELTGPKVLACSLA 121
DB 246 GGLHDSLSYRWVSPALPKDFLPQAKYQIRRVED--SYDMKVVDDVSNQ-----SCLRLA 298

RESULT 34
US-08-348-353-17

? Sequence 17, Application US/08348353
? Patent No. 5932217
? GENERAL INFORMATION:
? APPLICANT: Tuomenen, Elaine
? APPLICANT: Hunter, Robert
? APPLICANT: KIM, Anthony
? TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
? NUMBER OF SEQUENCES: 38
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Klauber & Jackson
? STREET: 411 Hackensack Avenue
? CITY: Hackensack
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07601
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/348.353
? FILING DATE: 30-NOV-1994
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: JACKSON, David 76.742
? REGISTRATION NUMBER: 1000-1-09/CIP1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-487-5800
? TELEFAX: 201-343-1684
? INFORMATION FOR SEQ ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1248
? TYPE: amino acids
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-348-353-17

Query Match 9.3%, Score 61.5, DB 2, Length 1248;
Best Local Similarity 30.6%, Pred No. 49;
Matches 22, Conservative 13, Mismatches 28, Indels 9, Gaps 3;
OY 26 FETQSLMASSHKL--TGGCAGDVGVGPPDSRLP-----VQENC---ADRPVHDSRA 76
DB 879 EAAIIMAGELITVKAQNTNNRAILMAGNABLYAAVALMLKGRIRACBMDHAPR 938
OY 77 TKHPLTGTQDQ 88
DB 939 IENFAKLSGQ 950

RESULT 35
US-08-465-965-17
? Sequence 17, Application US/08465965
? Patent No. 5968512
? GENERAL INFORMATION:
? APPLICANT: Tuomenen, Elaine
? APPLICANT: Hunter, Robert
? TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
? NUMBER OF SEQUENCES: 38
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Klauber & Jackson
? STREET: 411 Hackensack Avenue
? CITY: Hackensack
? STATE: New Jersey
? COUNTRY: U.S.A.
? ZIP: 07601
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible


```

CURRENT APPLICATION DATA:
US-08-439-009A-4
APPLICATION NUMBER: US/07/642,734C
CLASSIFICATION: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952, US-01
TELEPHONE: 08-37-93396
TELEPHONE: 708-38-65396
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

```

```

Query Match
Best Local Similarity 26.7%; Score 61; DB 2; Length 3567;
Pred. No. 2,4e+02;
Matches 31; Conservative 10; Mismatches 59; Indels 16; Gaps 4;

```

```

OY 14 TWCP-----VTEAIFETQXSLVA--ESBIXIKTLQCCDADVPGRPSRLPAVQENGA 66
DB 1127 SMPPIGVTVLGAASVQDDLVRLMDRGAERVLVAGAC-----PDDLLAAVEEGA 1179
OY 67 GEPVILDSPAIKIQPLTGGTCGRYSCISLSTGKXSLKLELGRKYVLAASIAL 122
DB 1180 SAVCQADAAALREAL--GDEPVYALVHAGTLYNFGSISFVAPEERETIAAKTAL 1233

```

```

RESULT 38
US-08-439-009A-4

```

```

Sequence 4; Application US/0843909A
Patent No. 6004787

```

```

GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Kcaipine, J B
APPLICANT: Kcaipine, J B
TITLE OF INVENTION: Method of directing biosynthesis of
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Steven F. Weinstein
STREET: Abbott Laboratories D377/AF6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patient Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952, US-01
TELEPHONE: 847-938-3137
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

US-08-439-009A-4

```

```

Query Match
Best Local Similarity 26.7%; Score 61; DB 3; Length 3567;
Pred. No. 2,4e+02;
Matches 31; Conservative 10; Mismatches 59; Indels 16; Gaps 4;

```

```

OY 14 TWCP-----VTEAIFETQXSLVA--ESBIXIKTLQCCDADVPGRPSRLPAVQENGA 66
DB 1127 SMPPIGVTVLGAASVQDDLVRLMDRGAERVLVAGAC-----PDDLLAAVEEGA 1179
OY 67 GEPVILDSPAIKIQPLTGGTCGRYSCISLSTGKXSLKLELGRKYVLAASIAL 122
DB 1180 SAVCQADAAALREAL--GDEPVYALVHAGTLYNFGSISFVAPEERETIAAKTAL 1233

```

```

RESULT 39
US-09-071-224-29

```

```

Sequence 29; Application US/09071224
Patent No. 6,231,313

```

```

GENERAL INFORMATION:
APPLICANT: Lok, SI
APPLICANT: Presnell, Scott R.
APPLICANT: Jelmeiry, Anna C.
APPLICANT: Gilbert, Teresa
APPLICANT: Foster, Donald C.
APPLICANT: Adams, Robyn L.
TITLE OF INVENTION: RAMMULIAN ZCYTORS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: Zymogenetics
STREET: 1201 Eastlake Ave East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastISO for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: NUMBER: US/09/071,224
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, S
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: 96-22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6637
TELEFAX: 206-442-6678
TELEFAX: 206-442-6678

```

```

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-224-29

```

```

Query Match
Best Local Similarity 26.7%; Score 60.5; DB 4; Length 389;
Pred. No. 13;
Matches 32; Conservative 10; Mismatches 57; Indels 21; Gaps 5;

```

```

OY 14 TWCP-----EAAIFETQXSLVAESBIXI-----KTLQCCDADVPGRPSRLPAV 61
DB 1127 SMPPIGVTVLGAASVQDDLVRLMDRGAERVLVAGAC-----PDDLLAAVEEGA 1179

```


Db 152 TYGPHSCHIPDIALF--TPVEIWEATNLSGASRVLTLLDITVTDPPEPVHVSV 209
 Oy 62 QEMGADPEVHIDPAIKHOFILITGDTQGRKRSGLSTGKXOLSLLEITGPVYACSLIA 121
 Db 210 GGLLEDSIVKVSPPALADPLFLOAKYOLIRVED--SDVMKVVDDVSNPT-----SCHLA 262

RESULT 40
 US-08-449-645A-19

Sequence 19, Application US/08449645A
 Patent No. 5981245

GENERAL INFORMATION:

APPLICANT: Fox, Gary M.

INVENTOR: Fox, Gary M. EPH-Like Receptor Protein Tyrosine

TITLE OF INVENTION: Kinases

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSER: Amgen Patent Operations/RBW

STREET: 1840 Dehavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/449,645A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-287

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-449-645A-19

Query Match 9.2% Score 60.5; DB 2: Length 984;
 Best Local Similarity 23.1%; Pred. No. 48;
 Matches 31: Conservative 13; Mismatches 45; Indels 45; Gaps 6;

Oy 17 PVEAIAIFETQXSLWA-----ESEHXIKTLGQCDADV---KGP 52
 Db 856 PVDCPAPLVEIKMCMAYDARRRPHOXLAHLEQLAMPISLRTIANPDPVYTLPLSL 915
 Oy 53 PQDSRLP--AVQMGADPEVHIDPAIKHOFILITGDTQGRKRSGLSTGKXOLSLLEI 110
 Db 916 SGGSDIPIYRTVSEW-----LESIMKRYTL-----HFSAGIDT---MECVLEL 956

Oy 111 TGPVYACSLIADG 124
 Db 957 TADLTQMGITLPG 970

RESULT 41
 US-08-702-367A-19
 Sequence 19, Application US/08702367A
 Patent No. 5981246
 GENERAL INFORMATION:
 APPLICANT: Fox, Gary M.
 INVENTOR: Fox, Gary M. EPH-Like Receptor Protein Tyrosine
 TITLE OF INVENTION: Kinases
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Amgen Patent Operations/RBW

STREET: 1840 Dehavilland Drive
 CITY: Thousand Oaks
 STATE: California
 COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,367A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-287

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-702-367A-19

Query Match 9.2% Score 60.5; DB 2: Length 984;
 Best Local Similarity 23.1%; Pred. No. 48;
 Matches 31: Conservative 13; Mismatches 45; Indels 45; Gaps 6;

Oy 17 PVEAIAIFETQXSLWA-----ESEHXIKTLGQCDADV---KGP 52
 Db 856 PVDCPAPLVEIKMCMAYDARRRPHOXLAHLEQLAMPISLRTIANPDPVYTLPLSL 915
 Oy 53 PQDSRLP--AVQMGADPEVHIDPAIKHOFILITGDTQGRKRSGLSTGKXOLSLLEI 110
 Db 916 SGGSDIPIYRTVSEW-----LESIMKRYTL-----HFSAGIDT---MECVLEL 956

Oy 111 TGPVYACSLIADG 124
 Db 957 TADLTQMGITLPG 970

RESULT 42
 PCT-US95-04681-19
 Sequence 19, Application PC/TUS9504681
 GENERAL INFORMATION:
 APPLICANT: Fox, Gary M.
 TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
 INVENTOR: Fox, Gary M.
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Amgen Patent Operations/RBW
 STREET: 1840 Dehavilland Drive
 CITY: Thousand Oaks
 STATE: California
 COUNTRY: USA
 ZIP: 91320
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/04681
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Winter, Robert B.
 REFERENCE/DOCKET NUMBER: A-287
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:

LENGTH: 984 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 MOLECULE TYPE: linear
 MOLECULE TYPE: protein
 PCT-US95-04661-19

Query Match 9.2%; Score 60.5; DB 5; Length 984;
 Best Local Similarity 23.1%; Pred. No. 48;

Matches 31; Conservative 13; Mismatches 45; Indels 45; Gaps 6;

17 PUTGAAIPETQXSLAA-----SEBXKTLGQCADV---GCP 52
 DB 866 PDCAPATLTKMKCAVDKRRPHRQKLAHLEQLANPISRTIANFDPRVTLPLSL 915
 QY 53 PDCSLP--AVQEMGADSPVHIDSPAIKHQFLITGDTQGRKCRSGLSGMXOLSLLEI 110
 DB 916 SSGSDIPEYFVTSER-----LESIMKRYTL-----HFSAGLDT---HECVLEL 956
 QY 111 TQRYLACSLALG 124
 DB 957 TARDITQWQITLPG 970

RESULT 43

US-08-667-939A-2
 Sequence 2, Application US/0867939A
 Patent No. 599165

GENERAL INFORMATION:

APPLICANT: LHO, Shun

TITLE OF INVENTION: CD16-IT VARIANTS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROMDY AND NEIMARK

STREET: 419 South Main Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

21P: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/667,939A

FILING DATE: 24-JUN-1996

PRIOR APPLICATION DATA:

FILING DATE: NUMBER: US 08/433,123

ATTORNEY/AGENT INFORMATION:

NAME: BROMDY, Roger L.

REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: LHO 2A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-628-5198

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS: 2:

LENGTH: 254 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-667-939A-2

Query Match 9.1%; Score 60; DB 2; Length 254;

Best Local Similarity 28.3%; Pred. No. 8.5;

Matches 34; Conservative 13; Mismatches 43; Indels 30; Gaps 7;

QY 3 MLVPELLMGVYMGVPE-----AAIEFTQ-SLMAESHKIKTLGQCADVPGPGISGR 57

DB 4 LLLPALLVLSAGMKRTEDLPAAVYLEPQWVSYLEKDSVTLKQCAVY-----PEDNS 57
 QY 58 LPAVQENGADEPHIDSPAIKHQFL--LTGDTQGRKCRSGLSF-----GMXOL 104
 DB 58 ----TQMFKE--NLISQSSFTIANTYDSSGTFQGMHSTLSDPQGLEVQVGMALL 111

RESULT 44

US-08-759-628-5

Sequence 5, Application US/08759628

Patent No. 623446

GENERAL INFORMATION:

APPLICANT: Altman, Scott W.

APPLICANT: Bazan, J. Fernando

APPLICANT: Kastelein, Robert A.

TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNA Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

21P: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/759,628

FILING DATE: 05-DEC-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

FILING DATE: NUMBER: US 60/008,574

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,050

REFERENCE/DOCKET NUMBER: DX0532Q

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-496-1200

TELEFAX: 415-496-1200

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 421 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-759-628-5

Query Match 9.1%; Score 60; DB 4; Length 421;

Best Local Similarity 23.7%; Pred. No. 17;

Matches 40; Conservative 14; Mismatches 59; Indels 56; Gaps 9;

QY 4 LVPFLIMQVYMGVPEYFAIIPYETQXSLMAESHKIKTLGQCADV-----GPP----- 53

DB 59 WFDDELPTVALEVVATNNPDKRVKPEFDQGRKRLADVQAKGSLSDGAKHEDTGSYFFR 118

QY 93 CRSGLSGKXG--QSLKTL--LELQ-----PVLASCL 120

DB 119 VENGKGVKYSQOKNKLNEVTLIERPDIHLGSPLESGLMFPRLSGSL 167

RESULT 45

```

US-07-683-957B-3
: Sequence 3, Application US/07683957B
: Patent No. 5310880
: GENERAL INFORMATION:
: INVENTOR: Goldstein, Patricia K.
: APPLICANT: Regio, Richard C.
: TITLE OF INVENTION: Purification of M llerian Inhibiting
: TITLE OF INVENTION: Substance
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
: FIRM: 1100 New York Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: DOS/MSDOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/683,957B
: FILING DATE: 19910412
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldstein, Patricia K.
: REGISTRATION NUMBER: 0609,3060000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 466-0800
: TELEFAX: (202) 833-8716
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: TYPE: amino acids
: TOPOLOGY: Linear
: MOLECULE TYPE: protein
US-07-683-957B-3

```

```

Query Match          9 18      Score 60: DB 1; Length 553;
Similarity 27.88;  Freq NO. 25;
Matches 32; Conservative 9; Mismatches 52; Indels 22; Gaps 6;

OY 20 EAATFEQXSLWASESHVKTGLQGDADVPQPGDSRIPAVO-----EMGAOE--PV 70
DB 77 EQATLELVQESHWGPOD--LATFGVCTSD-----SQTLPALQRIAGMLGFTGROQLIVL 129
OY 71 HLDSPAIKHQFLVLTGDTQIRNCRSGSLSTGAWO--LSKLELTPGKPVLAGSLALDG 124
DB 130 HLAETVETQLLAKFOEP-----PFGASRWQALVLYLPGFQVYTVTGAGLDG 179

```

Search completed: January 7, 2002, 16:49:55
Job time: 282 sec

•
•
•

•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

ON protein - protein search, using sw model

Run on: January 7, 2002, 16:44:08 ; Search time 31.99 seconds

(without alignments) 291.755 Million cell updates/sec

Title: 05-09-471-276-831

Perfect score: 1 KSHLVFLLMLWYMGVYE.....LLEITQRYVACSLADQAS 126

Sequence:

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database: A.Geneseq.1101.*

1: /SID52/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Accession	Length	DB ID	Description
1	655	59.1	126	21	AA64670
2	139.5	21.1	512	22	AA64874
3	95.5	14.4	299	22	AAU01393
4	92.5	14.4	299	22	AAU01394
5	91.5	13.8	265	21	AA604182
6	91.5	13.8	265	21	AA604183
7	91.5	13.8	299	22	AAU01390
8	91.5	13.8	299	22	AAU01392
9	91.5	13.8	299	22	AAU01392
10	90.5	13.7	623	19	AA669233
11	90.5	13.7	631	19	AA682552

12	90.5	13.7	631	21	AA604177	Leukocyte immunog1
13	87.5	13.2	631	21	AA61263	Human monocytic inh
14	85.5	12.9	447	19	AA69234	FeR-IV protein seq
15	84.5	12.8	478	19	AA62551	Human LIR-pbm2 pro
16	84.5	12.8	448	19	AA62551	Human gp48 mAb po
17	81.5	12.3	500	19	AA62553	Human LIR-pbm5 p
18	81.5	12.3	500	19	AA62553	Human LIR-pbm5 p
19	78.5	11.9	652	19	AA604178	Leukocyte immunog1
20	78.5	11.9	652	19	AA604178	Human LIR-183 pro
21	78.5	11.9	652	21	AA604170	Leukocyte immunog1
22	75.5	11.4	635	21	AA657323	Human LIR-pbm2 p
23	75.5	11.4	635	20	AA62550	Nos3c Nsp1 restr1
24	75.5	11.4	635	20	AA62550	Mouse dephosphoryl
25	74.5	11.3	663	20	AA62550	Human protein sequ
26	74.5	11.3	663	20	AA62550	Mouse dephosphoryl
27	74.5	11.3	680	20	AA62550	Human LIR-pbm2 p
28	74	11.2	289	19	AA62548	Human LIR-pbm5-2
29	74	11.2	289	21	AA604173	Leukocyte immunog1
30	74	11.2	650	19	AA62544	Human LIR-p32 pro
31	74	11.2	650	21	AA604189	Leukocyte immunog1
32	74	11.2	651	19	AA602782	Protein encoded by
33	73	11.0	153	19	AA61050	Chicken matrix met
34	73	11.0	153	19	AA61050	Chicken matrix met
35	73	11.0	193	19	AA61235	Alpha-v-beta-5 ant
36	73	11.0	128	19	AA61089	Chicken matrix met
37	73	11.0	228	19	AA61234	Alpha-v-beta-5 ant
38	73	11.0	663	19	AA61111	Chicken matrix met
39	73	11.0	663	19	AA61127	Chicken matrix met
40	72.5	11.0	244	21	AA62556	Nos3c sp restrict
41	72.5	11.0	244	21	AA62556	Human protein sequ
42	72.5	11.0	287	21	AA604181	Leukocyte immunog1
43	72.5	11.0	542	26	AA615104	Chimeric protein I
44	71.5	10.8	90	22	AA62555	Human protein sequ
45	71	10.7	324	12	AA612427	Hyb10 Fc(gamma)RI
46	70	10.6	191	22	AA69300	Human LIR-pbm3-4
47	69.5	10.5	341	22	AA65833	Murine mature IWE
48	69.5	10.5	370	22	AA65830	Human secreted pro
49	69.5	10.5	370	22	AA65830	Human secreted pro
50	69.5	10.5	370	22	AA65830	Human secreted pro
51	69.5	10.5	841	20	AA624318	Human platelet mem
52	69	10.4	339	22	AA627990	Human LIR-pbm3-4
53	69	10.4	489	19	AA682549	Leukocyte immunog1
54	69	10.4	489	21	AA604174	Human LIR-pbm3-4
55	68.5	10.4	244	21	AA62553	Arabinoside chilla
56	68.5	10.4	325	21	AA659024	Human LIR-pbm3-4
57	68.5	10.4	389	21	AA676303	Human LIR-pbm3-4
58	68.5	10.4	389	21	AA676303	Human LIR-pbm3-4
59	68.5	10.4	390	20	AA627096	Human LIR-pbm3-4
60	68.5	10.4	390	20	AA627096	Human LIR-pbm3-4
61	68.5	10.4	390	20	AA627096	Human LIR-pbm3-4
62	68.5	10.4	390	21	AA688574	Human secreted pro
63	68.5	10.4	390	21	AA688574	Human secreted pro
64	68.5	10.4	390	21	AA688574	Human secreted pro
65	68.5	10.4	390	22	AA606610	Human secreted pro
66	68.5	10.4	390	22	AA606610	Human secreted pro
67	68.5	10.4	390	22	AA606610	Human secreted pro
68	68.5	10.4	390	22	AA688599	Human membrane or
69	68.5	10.4	390	22	AA688599	Human membrane or
70	68.5	10.4	390	22	AA688599	Human membrane or
71	68.5	10.4	390	22	AA688599	Human membrane or
72	68.5	10.4	390	22	AA688599	Human membrane or
73	68.5	10.4	390	22	AA688599	Human membrane or
74	67.5	10.2	119	21	AA659188	Human LIR-pbm3-4
75	67.5	10.2	141	22	AA639287	Human LIR-pbm3-4
76	67.5	10.2	145	22	AA641053	Human LIR-pbm3-4
77	67.5	10.2	217	22	AA658871	Human LIR-pbm3-4
78	67.5	10.2	220	22	AA658871	Human LIR-pbm3-4
79	67.5	10.2	220	22	AA658871	Human LIR-pbm3-4
80	67.5	10.2	370	22	AA65906	Human secreted pro
81	67	10.1	16	22	AA615665	Schizophrenia-asso
82	67	10.1	16	22	AA615665	Schizophrenia-asso
83	67	10.1	16	22	AA615665	Schizophrenia-asso
84	67	10.1	16	22	AA687241	Breast-cancer asso

PX	
XX	M0200121631-A2.
PN	
XD	29-MAR-2001.
XX	
PD	20-SEP-2000; 2000M0-US25982.
PF	
XX	20-SEP-1999; 590US-0389723.
PR	
XX	(MULTL-) MILLENNIUM PHARM INC.
PA	
XI	Kirst SJ, Sharp JD, Fraser CC, Barnes T, Kingsbury G;
D1	WPI: 2001-211461/21.
DR	N-PDSB: AAS02104.
XX	
PT	New nucleic acid encoding INTERCEPT 307, MANCO 511, TANCO 351, TANCO
PT	361, TANCO 499 or TANCO 509 secreted or transmembrane protein, useful
PT	for the diagnosis and treatment of arthritis, psoriasis and Parkinson's
XX	disease -
XX	
PS	Disclosure: Page 299-300; 362pp; English.
XX	

Query Match	14.4%	Score 95.5	DB 22	Length 299
Best Local Similarity	27.0%	Pred. No. 0.0039		
Matches	34	Conservative 17	Mismatches 48	Indels 27
Gaps	5			
QY	1 MSNLYVLLKQYNGYNGPVEALFAEFLYKQSSIAHE-----SRHKLTGQGNQD-----	48		
DB	19 GSPALMVLLGILGIPHTVNGSKSLKTLWEGGVSISGNAVLCIGGLLEAGYRL	76		
QY	49 -VEPPQDSLEPAVQKCAQEPVHLIDSPKIRHQLFLTGQDQ--GNYKCRSLSTQKXL	104		
DB	79 VKEGSP-----EPWDTGNPL--EPKHKARFISPMLEHAGRYCYYSPPAGWSEP	127		
QY	105 SKLDEL	110		
DB	128 adpdel	133		
RESULT 4				
AA001394				
AA001394 standard: Protein: 299 AA.				
XX	AA001394:			
XX	18-JUL-2001 (first entry)			
XX				
DE	Human MANCO 511, variant #4 amino acid sequence.			
XX				
KW	Human; MANCO 511; transmembrane protein; adenosine; asthma;			
KW	immunological disorder; arthritis; graft rejection; renal disorder;			
KW	acquired immunodeficiency syndrome; inflammatory disorders; psoriasis;			
KW	prostate; cerebrovascular disease; pituitary; Cushing's disease;			
KW	neurodegenerative disease; Parkinson's disease.			
XX				
OS	Homo sapiens.			
XX				

PN WO200121631-A2.
 XX 29-MAR-2001.
 PD 20-SEP-2000: 2000MO-US25982.
 XX 20-SEP-1999: 99US-0399723.
 PR (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Kirtac SJ, Sharp JD, Fraser CC, Barnes T, Kingsbury G.
 PR MPI: 2001-21164/21.
 DR N-PSDB: AAS02105.
 XX New nucleic acid encoding INTERCEPT 307, MANCO 511, TANGO 351, TANGO 361, TANGO 499 or TANGO 509 secreted or transmembrane protein, useful for the diagnosis and treatment of arthritis, psoriasis and Parkinson's disease.
 PS Disclosure: Page 301-302; 362pp; English.
 CC The sequence represents the amino acid sequence of human MANCO 511 variant 14 transmembrane protein. The nucleic acid and polypeptide sequences are useful for the diagnosis, prognosis and treatment of immunological disorders (e.g. arthritis, graft rejection and acquired immunodeficiency syndrome), embryonic disorders (e.g. porphyria and asthma), renal disorders, cerebrovascular diseases (e.g. lacunae), CC tumours, prostate-related disorders, pituitary-related disorders (e.g. Cushing's disease) and neurodegenerative diseases (e.g. Parkinson's disease).
 CC
 SO Sequence 299 AA:
 Query Match 14.0%; Score 92.5; DB 22; Length 299;
 Best Local Similarity 26.2%; Pred. No. 0.0088;
 Matches 33; Conservative 18; Mismatches 48; Indels 27; Gaps 5;
 Oy 1 MSALVFLILMGVYTWGPTVEAIFETQXSLNAG-----SEHXLKTLQCDNAD---- 48
 Db 19 vspalmvllciglsiprlhwaglnsktlwdegsvlsrgnsvclrcqgleaqeyrl 78
 Oy 49 -VPGPQDSRLPANOEGKQPEVHLDSFAIKHDFLRTGPTG--GRYRCSGLSTGKXOL 104
 Db 79 vkegsp-----epwdtqpl---epkhkarfslpsmtehagryrcyyyspgwsep 127
 Oy 105 SKLLEL 110
 Db 128 sdplel 133
 RESULT 5
 AAB04182 standard; Protein: 265 AA.
 XX AAB04182;
 XX 11-APR-2001 (first entry)
 DE Leukocyte immunoglobulin like receptor LIR-9s1.
 XX Leukocyte immunoglobulin like receptor; LIR; gene therapy;
 KW autoimmunity; autoimmune disorders; immune system; human.
 OS Homo sapiens.
 XX WO200068383-A2.
 PN 16-NOV-2000.
 PD 12-MAY-2000: 2000MO-US13228.
 DR

XX 12-MAY-1999: 99US-0310463.
 PR (IMV) IMMUNEX CORP.
 XX Cosman DJ, Anderson DM, Borges L;
 PR MPI: 2000-687645/67.
 DR N-PSDB: AAS4510.
 XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for treating autoimmune diseases and disease states with suppressed immune function.
 PS Claim 1: Page 111; 117pp; English.
 CC Host cells transformed with a vector which are capable of expressing a leukocyte immunoglobulin like receptor (LIR), can be used to produce the LIR polypeptide. LIR coding sequences may be used in the gene therapy of disorders mediated directly or indirectly by effects of the LIR polypeptides. The LIR polypeptides can be used to treat autoimmune diseases and disease states with suppressed immune function.
 CC
 SO Sequence 265 AA:
 Query Match 13.8%; Score 91.5; DB 21; Length 265;
 Best Local Similarity 26.2%; Pred No. 0.0099;
 Matches 33; Conservative 18; Mismatches 48; Indels 27; Gaps 5;
 Oy 1 MSALVFLILMGVYTWGPTVEAIFETQXSLNAG-----SEHXLKTLQCDNAD---- 48
 Db 19 vspalmvllciglsiprlhwaglnsktlwdegsvlsrgnsvclrcqgleaqeyrl 78
 Oy 49 -VPGPQDSRLPANOEGKQPEVHLDSFAIKHDFLRTGPTG--GRYRCSGLSTGKXOL 104
 Db 79 vkegsp-----epwdtqpl---epkhkarfslpsmtehagryrcyyyspgwsep 127
 Oy 105 SKLLEL 110
 Db 128 sdplel 133
 RESULT 6
 AAB04180 standard; Protein: 299 AA.
 XX AAB04180;
 XX 11-APR-2001 (first entry)
 DE Leukocyte immunoglobulin like receptor LIR-9m1.
 XX Leukocyte immunoglobulin like receptor; LIR; gene therapy;
 KW autoimmunity; autoimmune disorders; immune system; human.
 OS Homo sapiens.
 XX WO200068383-A2.
 PN 16-NOV-2000.
 PD 12-MAY-2000: 2000MO-US13228.
 PR 12-MAY-1999: 99US-0310463.
 PR (IMV) IMMUNEX CORP.
 XX Cosman DJ, Anderson DM, Borges L;
 PR MPI: 2000-687645/67.
 DR N-PSDB: AAS4510.
 XX

XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
 PT treating autoimmune diseases and disease states with suppressed immune
 PT function
 XX
 XX Claim 1: Page 106, 117pp: English.
 CC Host cells transformed with a vector which are capable of
 CC expressing a leukocyte immunoglobulin-like receptor (LIR), can be
 CC used to produce the LIR polypeptide. LIR coding sequences may be
 CC used in the gene therapy of disorders mediated directly or
 CC indirectly by defective or insufficient amounts of any of the LIR
 CC polypeptides. The LIR polypeptides can be used to treat autoimmune
 CC diseases and disease states with suppressed immune function.
 XX
 XX Sequence 299 AA:

Query Match 13.8%; Score 91.5; DB 21; Length 299;
 Best Local Similarity 26.2%; Pred. No. 0.012; Indels 27; Caps 5;
 Matches 33; Conservative 18; Mismatches 48; Indels 27; Caps 5;

OY 1 MSMLVFLILGWTGMPVTEALIFETQXSLAE-----SEHIXLKTIGCCDAQ---- 48
 Db 19 VSPALMVLICIGISIPRTHVQANISKALVWEGSVISGNSVILRCGFLAEQGYI 78
 OY 49 -VPGPDSRLPAVQWKAQEPVHLDSPAIKHQIFLLGDTQ--GRRYCRSGSLSTGMQL 104
 Db 79 VKESGP-----EPKDCQNP1--EPKHAKISIPSMCHAGRYCYGYSPPSWEEP 127
 OY 105 SKLEL 110
 Db 128 SDPEL 133

RESULT 7
 ID AA001330 standard: Protein: 299 AA.
 XX AA001330:
 XX 18-JUL-2001 (first entry)
 XX Human MANCO 511 amino acid sequence.
 XX
 XX Human: MANCO 511: transmembrane protein; diagnostic: asthma;
 KW immunological disorder: arthritis; graft rejection; renal disorder;
 KW acquired immunodeficiency syndrome; inflammatory disorders; psoriasis;
 KW AIDS; embryonic disorder: brain; cerebral oedema; ischaemia; tumour;
 KW prostate; cerebrovascular disease; pituitary; Cushing's disease;
 KW neurodegenerative disease: Parkinson's disease.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PT Peptide 1..41
 PT Protein 42..299
 PT Domain 42..265 "Mature MANCO 511"
 PT 43..46 "Extracellular domain"
 PT Modified-site 43..46 /note- "Asn is N-glycosylated"
 PT Modified-site 54..59 /note- "N-myristylation site"
 PT Modified-site 60..65 /note- "N-myristylation site"
 PT Domain 60..118 "N-myristylation site"
 PT 60..118 "Immunoglobulin (Ig)-like domain"
 PT Modified-site 64..66 /note- "Protein kinase C phosphorylation site"
 PT Modified-site 69..74 /note- "N-myristylation site"

PT Modified-site 105..108 /note- "Casein kinase II phosphorylation site"
 PT Modified-site 153..158 /note- "Casein kinase II phosphorylation site"
 PT Modified-site 157..160 /note- "Asn is N-glycosylated"
 PT Modified-site 193..198 /note- "N-myristylation site"
 PT Modified-site 207..209 /note- "Protein kinase C phosphorylation site"
 PT Modified-site 217..219 /note- "Protein kinase C phosphorylation site"
 PT Modified-site 228..232 /note- "N-myristylation site"
 PT Modified-site 248..251 /note- "Asn is N-glycosylated"
 PT Domain 266..284 /note- "Transmembrane domain"
 PT Domain 285..299 /note- "Cytoplasmic domain"
 PT Modified-site 289..299 /note- "Protein kinase C phosphorylation site"

MO200121631-A2.
 XX 29-MAR-2001.
 XX 20-SEP-2000: 2000MO-US25982.
 XX 20-SEP-1999: 990US-0399723.
 XX (MILL-) MILLERNIUM PHARM INC.
 XX
 XX Kirst SJ, Fraser CC, Barnes T, Kingsbury G:
 XX WPI: 2001-211451/21.
 XX DR N-PSDB; NAs02005.
 XX New nucleic acid encoding INTERCEPT 307, MANCO 511, TANGO 351,
 PT 351, TANGO 459 or TANGO 509 secreted or transmembrane protein, useful
 PT for the diagnosis and treatment of arthritis, psoriasis and Parkinson's
 PT disease -
 XX
 XX Claim 8: Fig 7; 362pp: English.
 XX
 XX The sequence represents the amino acid sequence of human MANCO 511
 CC transmembrane protein. The nucleic acid and polypeptide sequences
 CC are useful for the diagnosis, prognosis and treatment of immunological
 CC disorders (e.g. arthritis, graft rejection and acquired immunodeficiency
 CC syndrome), inflammatory disorders (e.g. psoriasis and asthma), renal
 CC disorders, embryonic disorders, brain-related disorders (e.g. cerebral
 CC oedema), cerebrovascular diseases (e.g. ischaemia), tumours, prostate
 CC cancer, neurodegenerative diseases (e.g. Parkinson's disease),
 CC and neurodegenerative diseases (e.g. Parkinson's disease).
 XX
 XX Sequence 299 AA:

Query Match 13.8%; Score 91.5; DB 22; Length 299;
 Best Local Similarity 26.2%; Pred. No. 0.012; Indels 27; Caps 5;
 Matches 33; Conservative 18; Mismatches 48; Indels 27; Caps 5;

OY 1 MSMLVFLILGWTGMPVTEALIFETQXSLAE-----SEHIXLKTIGCCDAQ---- 48
 Db 19 VSPALMVLICIGISIPRTHVQANISKALVWEGSVISGNSVILRCGFLAEQGYI 78
 OY 49 -VPGPDSRLPAVQWKAQEPVHLDSPAIKHQIFLLGDTQ--GRRYCRSGSLSTGMQL 104
 Db 79 VKESGP-----EPKDCQNP1--EPKHAKISIPSMCHAGRYCYGYSPPSWEEP 127
 OY 105 SKLEL 110
 Db 128 SDPEL 133

RESULT 8

AAU01391
ID AAU01391 standard; Protein: 299 AA.

AC AU013917

DT 18-JUL-2001 (first entry)

Human MANCO 511, variant #1 amino acid sequence

XX Human: MANGO 511; transmembrane protein; diagnostic: asthma;
 KM immunological disorder; atrialitis; graft rejection; renal disorder;
 KM acquired immunodeficiency syndrome; inflammatory disorders; porphyria;
 KM AIDS; embryonic disorder; brain; cerebral oedema; ischaemia; tumour;
 KM prostatic; cerebrovascular disease; pituitary; Cushing's disease;
 KM neurodegenerative disease; Parkinson's disease.

05 Homo sapiens.

PN WO200121631-A2

PD 29-MAR-2001

PF 20-SEP-2000; 2000WO-US25982.

PR 20-SEP-1999; 99US-0399723.

PA (MILL-) MILLENNIUM PHARM INC.

PI Klirst SJ, Sharp JD, Fraser CC, Barnes T, Kingsbury G;

DR WPL; 2001-211461/21.

Not available

PT disease -

PS Disclosure; Page 295-296; 362pp; English

XX The sequence represents the amino acid sequence of human MME0 511
CC variant #1 transmembrane protein. The nucleic acid and polypeptide
CC sequences are useful for the diagnosis, prognosis and treatment of
CC immunological disorders (e.g., arthritis, graft rejection and acquired
CC immunodeficiency).
CC (e.g., cerebral disorders, embryonic disorders (e.g., ischaemia), disorders
CC (asthma), renal disorders, cerebrovascular diseases (e.g., Cushing's
CC (e.g., cerebral oedema), cerebrovascular diseases (e.g., Parkinson's
CC tumours, prostate-related disorders, pituitary-related disorders (e.g.
CC Cushing's disease) and neurodegenerative diseases (e.g., Parkinson's
CC disease).

Sequence 299 AA:

Query Match 13.8%; Score 91.5; DB 22; Length 299

Matches	33;	Conservative	18;	Mismatches	48;	Indels	27;	Gaps	5
---------	-----	--------------	-----	------------	-----	--------	-----	------	---

[illegible]

RESULT

ID AAU01392 standard; Protein: 299 AA.

AC AAU01392

DT 18-JUL-2001 (first entry)

DE Human MANCO 511, variant #2 amino acid sequence.

XX Human: MANGO 511; Transmembrane protein; diagnostic: asthma;
 KM Immunological disorder: atrialitis; graft rejection; renal disorder;
 KM acquired immunodeficiency syndrome; inflammatory disorders; psoriasis;
 KM AIDS; embryonic disorder; brain; cerebral oedema; ischaemia; tumour;
 KM prostatic; cerebrovascular disease; pituitary; Cushing's disease;
 KM neurodegenerative disease; Parkinson's disease.

OS Homo sapiens.

PN WO200121631-A2

PD 29-MAR-2001

PF 20-SEP-2000; 2000WO-US25982.

PR 20-SEP-1999; 99US-0399723.

PA (MILL-) MILLENNIUM PHARM INC.

PI Kirst SJ, Sharp JD, Fraser CC, Barnes T, Klingsbury G;

DR WP1; 2001-211461/21

[illegible]

PT disease -

PS Disclosure; Page 297-298; 362pp; English.

XX The sequence represents the amino acid sequence of human MANGO 511
CC variant #2 transmembrane protein. The nucleic acid and polypeptide
CC sequences are useful for the diagnosis, prognosis and treatment of
CC immunological diseases (e.g., atrophic, graft rejection and acquired
CC immunodeficiency syndrome), autoimmune diseases (e.g., rheumatoid arthritis,
CC asthma, psoriasis), cerebrovascular disorders (e.g., stroke) and disorders
CC (e.g., cerebral oedema), cerebrovascular diseases (e.g., ischaemia),
CC tumours, prostate-related diseases, pituitary-related disorders (e.g.,
CC Cushing's disease) and neurodegenerative diseases (e.g., Parkinson's
CC disease).

SQ Sequence 299 AA;

Query Match 13.88; Score 91.5; DB 22; Length 299;

Matches	33;	Conservative	18;	Mismatches	48;	Indels	27;	Gaps	5;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

```

OY      1  NSMLVLLLLLMGVMGYNVEALAFPEFKQSLMAE-----SBHXIKTLDOGDAD---- 48
Db      19  VSPRLNVLLQGLGISLGRTHVWAGNLSKELWEEPGVSLVSGNVSLTCGLTLEAEYSL 78
OY      49  -VROPDSSKLDAVQEKGAQDEPNNLSDFALKIQLLTCTDQ-----GRYKNSGLSTQKKL 104
Db      79  KLGSP-----EPQDQPLI-----EPKIKETLSIPMLTGHNGTLYQYVSPAGVSEPR 127
OY      105  SKLESL 110
Db      128  scpieL 133

```

RESULT	10
AA669223	AA669223
ID	AA669223 standard; Protein: 623 AA.
XX	
XX	AA669223:
XX	
XX	20-OCT-1998 (first entry)
XX	
XX	PCR-III protein sequence.
DE	
XX	
XX	Fe receptor-like protein; phagocytosis inducer; rheumatoid arthritis;
XX	immune complex related disease; systemic lupus erythematosus; allergy;
KM	hemolytic anaemia; thrombocytopenia; anaphylaxis; cancer; lymphoma;
KM	leukaemia; infection; immunomodulator; viral entry inhibitor; therapy;
XX	PCR-III.
XX	
OS	Homo sapiens.
XX	
FX	Key
FX	Location/Qualifiers
FX	1..16
FX	/note="signal peptide"
FX	17..623
FX	/note="mature FcR-III"
FX	
FX	NO9831806-R2.
XX	
XX	23-JUL-1998.
XX	
XX	20-JAN-1998: 98MO-USO184.
XX	
XX	18-JUN-1997: 97US-0049872
XX	21-JAN-1997: 97US-0034205.
XX	
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	
XX	Genetz RL, Murphy M, NL J, Olsen HS, Ruben SM:
XX	NP1: 1998-414105/35.
XX	N-PSDB; AA04826.
XX	
XX	Nucleic acid encoding Fe receptor-like polypeptides or their
XX	fragments - and related vectors, transformed cells and antibodies,
XX	immune screening and diagnosing diseases of the hematopoietic
XX	and immune systems
XX	
XX	Claim 23: Fig 3a, 14pp: English.
XX	
XX	This sequence is the Fc receptor-like II protein (FcR-III) of the
XX	invention. Cells containing the DNA are used to express the recombinant
XX	protein, and to screen for specific (ant)agonists. The proteins are used
XX	to induce phagocytosis, and their (ant)agonists are used to treat immune
XX	complex related diseases (e.g., rheumatoid arthritis, systemic lupus
XX	erythematosus, hemolytic anaemia, thrombocytopenia, anaphylaxis,
XX	allergy, colorectal or breast cancer, lymphoma, leukaemia, cancer,
XX	intercellular pathogens etc). The antagonists are also useful as an
XX	immunomodulators and inhibitors of viral (e.g. human immune deficiency or
XX	denge viruses) entry into cells. The proteins may also be used to screen
XX	(Ab), for specific binding agents, i.e. (ant)agonists, for raising antibodies
XX	be used therapeutically as antagonists; as assay reagents for diagnostic
XX	determination of the levels of expression of the proteins and for
XX	screening of the levels of expression of the proteins and for
XX	useful in hybridisation probes or primers for isolation of genes,
XX	in situ hybridisation (chromosome mapping) and diagnostically to measure
XX	mRNA expression.
XX	
XX	Sequence 623 AA:

Query Match	Score	DB	Length
13.7%	90.5	19	623

Matches	39;	Conservative	10;	Mismatches	38;	Indels	51;	Gaps	6;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

[illegible]

(LIR) polypeptide LIR-pmb17. This sequence can be administered

—

therapeutically to treat disorders associated with insufficient/defective amounts of LIR polypeptide. LIR-P362 and certain other LIR family members contain cytoplasmic immunoreceptor tyrosine-based inhibitory motifs (ITIMs), whilst other LIR family members lack ITIMs. By analogy with the known function of ITIMs in inhibitory receptors, it is expected that LIRs having ITIMs are inhibitory receptors mediating negative signalling, whilst those lacking ITIMs are activatory receptors. Failure of a receptor that mediates negative signalling could result in autoimmune diseases, whilst failure of a receptor mediating activatory signalling could result in suppressed immune function. They are also useful to produce probes for detecting LIR nucleic acids or isolating LIR DNA from other species.

Sequence 631 AA:

Query Match 13.7%; Score 90.5; DB 19; Length 631;
Best Local Similarity 28.3%; Pred. No. 0.041;
Matches 39; Conservative 10; Mismatches 38; Indels 51; Gaps 6;

OY 1 MSNLVFLILKCTYKCPTEAIFETQXSLANSESHXKTKLCCDADVPFGDSRLPA 60

Db 1 mpeatlclcgslsgrtrvqagrpfpklvne-----pg-----s 37

OY 61 VDRMG-----AOE-----PVHLD-----SPAIRKQFLITGDTQ--GRRY 92

Db 38 vlsqspvrlwcqgslaqeqyldkqgspepldtrmpkxkfalslpmqlqhagryr 97

OY 93 CMSGSTGCKXOLSKLEI 110

Db 98 chyyasgysgsepsdpl 115

RESULT 12

AA604177 standard; Protein: 631 AA.

AA604177:

11-APR-2001 (first entry)

Leukocyte immunoglobulin like receptor plm17.

Leukocyte immunoglobulin like receptor; LIR; gene therapy;

autoimmunity; autoimmune disorders; immune system; human.

Homo sapiens.

WO2000068383-A2.

16-NOV-2000.

12-MAY-2000: 2000MO-US13228.

12-MAY-1999: 99US-0310463.

(IMV) IMUNEX CORP.

Cosman DJ, Anderson DM, Borges L:

WPI: 2000-687645/67.

N-PSDB: AAA54601.

Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for treating autoimmune diseases and disease states with suppressed immune function

Example 9: Page 95-97: 117pp: English.
Host cells transformed with a vector which are capable of expressing a leukocyte immunoglobulin like receptor (LIR), can be used to produce the LIR polypeptide. LIR coding sequences may be used in the gene therapy of disorders mediated directly or

Indirectly by defective or insufficient amounts of any of the LIR polypeptides. The LIR polypeptides can be used to treat autoimmune diseases and disease states with suppressed immune function.

Sequence 631 AA:

Query Match 13.7%; Score 90.5; DB 21; Length 631;
Best Local Similarity 28.3%; Pred. No. 0.041;
Matches 39; Conservative 10; Mismatches 38; Indels 51; Gaps 6;

OY 1 MSNLVFLILKCTYKCPTEAIFETQXSLANSESHXKTKLCCDADVPFGDSRLPA 60

Db 1 mpeatlclcgslsgrtrvqagrpfpklvne-----pg-----s 37

OY 61 VDRMG-----AOE-----PVHLD-----SPAIRKQFLITGDTQ--GRRY 92

Db 38 vlsqspvrlwcqgslaqeqyldkqgspepldtrmpkxkfalslpmqlqhagryr 97

OY 93 CMSGSTGCKXOLSKLEI 110

Db 98 chyyasgysgsepsdpl 115

RESULT 13

AA61263 standard; Protein: 631 AA.

AA61263:

04-APR-2001 (first entry)

Human monocyte inhibitory receptor precursor.

Human: monocyte inhibitory receptor precursor; TANKO 368; cardiant;

cerebroprotective; cytoprotective; anticoagulant; thrombolytic;

antiarteriosclerotic; haemostatic; glycoprotein VI; GPII;

platelet membrane glycoprotein receptor; bleeding disorder;

blood vessel injury; thrombotic disorder; haemorrhagic disorder;

ischemia; cardiovascular disease; immunological disease; liver disorder;

cancer.

Homo sapiens.

WO200100810-A1.

04-JAN-2001.

30-JUN-2000: 2000MO-US18152.

30-JUN-1999: 99US-0345468.

06-DEC-1999: 99US-0454824.

14-FEB-2000: 2000US-0503387.

(MILL-) MILLENNIUM PHARM INC.

Busfield SJ, Vallejal J, Jandrot-Perrus W, Valinchenker W, Gill DS:

Qian MD, Kingsbury G:

WPI: 2001-080877/09.

N-PSDB: AMF29472.

New genes encoding human platelet-expressed collagen receptor, glycoprotein VI, and its modulators, useful for preventing, treating and diagnosing hemorrhagic disorders, thrombotic diseases and immunological disorders

Disclosure: Fig 4A-4B: 227pp: English.
The present sequence is given in a specification relating to an isolated nucleic acid molecule encoding a platelet membrane glycoprotein receptor glycoprotein VI (GPIV), also called TANKO 268. The GPIV polynucleotides and polypeptides and their modulators, e.g. antisense nucleic acids,

[illegible]

```

XX 12-MAY-1999: 99US-0310463.
XX (IMM) ) IMMUNEX CORP.
XX
XX Cosman DJ, Anderson DM, Borges L:
XX WPI: 2000-687645/67.
XX N-PSDB: AAA54600.
XX
XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
XX treating autoimmune diseases and disease states with suppressed immune
XX function
XX
XX Example 9: Page 91-92: 117pp: English.
XX
XX Host cells transformed with a vector which are capable of
XX expressing a leukocyte immunoglobulin-like receptor (LIR), can be
XX used to produce the LIR polypeptide. LIR coding sequences may be
XX used in the gene therapy of disorders mediated directly or
XX indirectly by defective or insufficient amounts of any of the LIR
XX polypeptides. The LIR polypeptides may be used to treat autoimmune
XX diseases and disease states with suppressed immune function.
XX
XX Sequence 448 AA:
SQ
Query Match 12.8%; Score 94.5; DB 21: Length 448:
Best Local Similarity 27.9%; Fred. No. 0.13; 44: Indels 33: Gaps 5:
Matches 34: Conservative 11: Mismatches 44:
OY 8 LLLGCVTMCPTGAIAFYETQXSIAME-----SRHKLMTGGCDA-----DVPQ 51
DB 8 ILLGLISYPRILMGAGPLPKPLWEPGVSIVSGVSCGLQLEATEYFLDKESPP 67
OY 52 PRCDSRLPAVDKMGKQCEPHLSDPAIKRQ--LLTGDTGCRRCRSGSLSTGKQSLKL 108
DB 68 P-----VDTQPL--EPKHKARISJPMEDVDYAGYTCYTRYPGWGSPDQPL 113
OY 109 EL 110
DB 114 el 115
RESULT 18
AAB82553
ID AAB82553 standard: Protein: 590 AA.
XX
XX AAB82553:
XX
XX 09-FEB-1999 (first entry)
XX
XX Human LIR-phmew protein.
XX
XX LIR-phmew: immunoregulator; leukocyte immunoglobulin-like receptor;
XX therapeutic; treatment; disorder; ITIM; MHC class I receptor; inhibitor;
XX cytoplasmic immunoreceptor tyrosine-based inhibitory motif; activator;
XX negative signalling; autoimmune disease; suppressor; LIR.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..16
XX Domain 1..456
XX FT /label= extracellular_domain
XX FT 1..590
XX FT /label= LIR-phmew
XX FT 457..579
XX FT /label= transmembrane_domain
XX FT 580..590
XX FT /label= cytoplasmic_domain
XX FT 554..557
XX Region

```

```

XX Region /note="ITIM motif"
XX 584..587
XX /note="ITIM motif"
XX
XX W09848017-A1.
XX
XX 29-OCT-1998.
XX
XX 23-APR-1998: 98WO-0508244.
XX
XX 24-APR-1997: 97US-0842248.
XX
XX (IMM) ) IMMUNEX CORP.
XX
XX Cosman DJ.
XX
XX WPI: 1998-605990/51.
XX N-PSDB: AAB69239.
XX
XX Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful,
XX e.g. for treating autoimmune diseases or disease states associated
XX with suppressed immune function
XX
XX Claim 4: Page 94-96: 112pp: English.
XX
XX This sequence represents a novel leukocyte immunoglobulin-like receptor
XX (LIR) polypeptide LIR-phmew. This sequence can be administered
XX therapeutically to treat disorders associated with insufficient/defective
XX amounts of LIR polypeptide. LIR-P302 and certain other LIR family members
XX (LITIMs). Whilst other LIR family members lack ITIMs by analogy with the
XX structure and function of known MHC class I receptor molecules, LIRs
XX having ITIMs are inhibitory receptors mediating negative signalling.
XX Whilst those lacking ITIMs are activatory receptors. Failure of a
XX receptor that mediates negative signalling could result in autoimmune
XX diseases. Whilst failure of a receptor mediating activatory signalling
XX could result in suppressed immune function. This sequence is used to
XX produce probes for detecting LIR nucleic acids or isolating LIR DNA from
XX other species.
XX
XX Sequence 590 AA:
SQ
Query Match 12.3%; Score 81.5; DB 19: Length 590:
Best Local Similarity 27.9%; Fred. No. 0.14; 44: Indels 27: Gaps 6:
Matches 38: Conservative 16: Mismatches 55:
OY 1 MSLLVPELLGCVTMCPTGAIAFYETQXSIAMESRHXIK-----TL--GGCDA-----47
DB 1 MLTSLVLCIGLSYPRILMGAGPLPKPLWEPGVSIVSGVSCGLQLEATEYFLDKESPP 60
OY 48 DVEKPCDSRLPAVDKMGKQCEPHLSDPAIKRQPLLTGDTGCRRCRSGSLSTGKQSLKL 107
DB 61 dkegiparkrqnprleghakfhiastv-----dsagrytcytrypgwspsdpd 112
OY 108 LEL--TG---PKVLA 117
DB 113 lclvalglfaepcllla 128
RESULT 19
AAB04178
ID AAB04178 standard: Protein: 590 AA.
XX
XX AAB04178:
XX
XX 11-APR-2001 (first entry)
XX
XX Leukocyte immunoglobulin like receptor phmew.
XX
XX Leukocyte immunoglobulin like receptor; LIR; gene therapy;
XX autoimmunity; autoimmune disorders; immune system; human.
XX
XX

```


CC of malignant tumours, haemophaly, HIV infection, immunological diseases
CC various inflammatory diseases. The present sequence is the
CC polypeptide of the invention.

XX Sequence 647 AA:

Query Match 11.3%: Score 75; DB 22; Length 647;

Best Local Similarity 31.5%: Pred. No. 3;

Matches 29; Conservative 10; Mismatches 23; Indels 30; Gaps 6;

OY 33 ASSEKMLTKLCCDADVRGRC-----DSRLPAVDGKQEPVHLDSF-AIRHQ 80

DB 192 addeavlnlneceddgypppsstfsqkqreadr-----gsqesnhllspqlyqyq 244

OY 81 FLITGDT-QGRYRCRSGD-----STGCMQL 104

DB 245 ---cgdsqkgyrctrrsrgcgclekmpwsl 273

RESULT 26

AAAY24322

ID AAY24322 standard; Protein: 663 AA.

XX AAY24322:

16-SEP-1999 (first entry)

Mouse dephosphorylase inhibiting p91-like protein #5.

Dephosphorylase inhibiting protein; p91: tyrosine phosphatase SHP-1;

SHP-2; inositol-5-phosphate SHP; phosphorylating tyrosine;

immunoreceptor; immunomodulatory agent.

Mus sp.

JP11169184-A.

29-JUN-1999.

12-DEC-1997: 97JP-0362285.

12-DEC-1997: 97JP-0362285.

(UYOK-) UNITV OKAYAMA.

WPI: 1999-422622/36.

N-PSDB: AAX88979.

New peptide - useful for inhibiting dephosphorylase

Claim 2; Page 22-24; 30pp; Japanese.

The present invention describes new proteins for inhibiting

dephosphorylase. The proteins can be combined with tyrosine phosphatase

SHP-1, SHP-2 or inositol-5-phosphatase SHP by phosphorylating tyrosine.

The present invention also describes an immunoreceptor comprising one

of the above proteins, and DNA coding the above proteins. The new proteins

can be used as an immunomodulatory agent. The present sequence

represents a protein from the present invention.

Sequence 663 AA:

DB 261 ksqktgtsqtlqpgnxygkflfipsvqghaggyrcyysasgwspsdclcl 312

RESULT 27

AAAY24320

ID AAY24320 standard; Protein: 680 AA.

XX AAY24320:

16-SEP-1999 (first entry)

Mouse dephosphorylase inhibiting p91-like protein #3.

Dephosphorylase inhibiting protein; p91: tyrosine phosphatase SHP-1;

SHP-2; inositol-5-phosphate SHP; phosphorylating tyrosine;

immunoreceptor; immunomodulatory agent.

Mus sp.

JP11169184-A.

29-JUN-1999.

12-DEC-1997: 97JP-0362285.

12-DEC-1997: 97JP-0362285.

(UYOK-) UNITV OKAYAMA.

WPI: 1999-422622/36.

N-PSDB: AAX88979.

New peptide - useful for inhibiting dephosphorylase

Claim 2; Page 18-19; 30pp; Japanese.

The present invention describes new proteins for inhibiting

dephosphorylase. The proteins can be combined with tyrosine phosphatase

SHP-1, SHP-2 or inositol-5-phosphatase SHP by phosphorylating tyrosine.

The present invention also describes an immunoreceptor comprising one

of the above proteins, and DNA coding the above proteins. The new proteins

can be used as an immunomodulatory agent. The present sequence

represents a protein from the present invention.

Sequence 680 AA:

Query Match 11.3%: Score 74.5; DB 20; Length 680;

Best Local Similarity 28.6%: Pred. No. 3.6;

Matches 32; Conservative 11; Mismatches 48; Indels 21; Gaps 4;

OY 6 VFLLMGVGMPTVE-----ATFEYKXSLMAESKMLTKLCCDADVRGPRGDSRLNAV 61

DB 215 vellsgnlgkprlkkepgsvltakamlwq-----gldaev-----yflhne 260

OY 62 QEWGQEPVHLDSFPAIKHQLTGDTO---GRYRCRSGLSTGXOLSKLLEL 110

DB 261 ksqktgtsqtlqpgnxygkflfipsvqghaggyrcyysasgwspsdclcl 312

RESULT 28

AAM82548

ID AAM82548 standard; Protein: 289 AA.

XX AAM82548:

09-FEB-1999 (first entry)

human LIR-pbm36-2 protein.

LIR-pbm3: Immunoregulator; leukocyte immunoglobulin-like receptor; LIR;

therapeutic; treatment; disorder; TTV: MHC class I receptor; inhibitor;

KW cytoplasmic immunoreceptor tyrosine-based inhibitory motif; activator;
 KW negative signalling; autoImmune disease; suppressor.

Homo sapiens.

Key Location/Qualifiers
 FT Peptide 1..16
 FT /label= signal
 FT Domain 1..261
 FT /label= extracellular_domain
 FT Protein 17..289 LIR-pmb36-2
 FT /label= LIR-pmb36-2
 FT Domain 281..289
 FT /label= transmembrane_domain
 FT Domain 281..289
 FT /label= cytoplasmic_domain

MO9848017-AL.

29-OCT-1998.

23-APR-1998: 98MO-US08244.

24-APR-1997: 97US-084224B.

(IMM) IMMUNEX CORP.

Cosman DJ.

WPI: 1998-60990/51.

N-PSDB: AA059334.

Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful,
 e.g. for creating autoimmune diseases or disease states associated
 with suppressed immune function

Claim 4: Page 73-74; 112pp; English.

This sequence represents a novel leukocyte immunoglobulin-like receptor
 (LIR) polypeptide LIR-pmb36-2. This sequence can be administered
 therapeutically to treat disorders associated with insufficient/defective
 amounts of LIR polypeptide. LIR-p362 and certain other LIR family members
 (LIRIM). Whilst other LIR family members lack ITIMs by analogy with the
 structure and function of known MHC Class I receptor molecules, LIRs
 having ITIMs are inhibitory receptors mediating negative signalling,
 whilst those lacking ITIMs are activatory receptors. Failure of a
 receptor that mediates negative signalling could result in autoimmune
 diseases, whilst failure of a receptor mediating activatory signalling
 could result in suppressed immune function. They are also useful to
 produce models for detecting LIR nucleic acids or isolating LIR DNA from
 other species.

Sequence 289 AA:

Query Match 11.2% Score 74; DB 19; Length 289;

Best Local Similarity 27.7% Pred. No. 1.4;

Matches 36; Conservative 15; Mismatches 45; Indels 34; Gaps 7;

1 MSMLVFLILGCVTGMPTGAIFETGXSIAMSESHXLTQ-----CDN----- 47

1 mprllvllcglstlprthvqagllprllwepgsavl-vgspvltvcgllletqeyr 59

48 ---DVPRPCDSRLPAVQENGAQEPYNIDSPAIKHOF---LITGTCGRYRCNSGLST-G 100

60 lyrekkcapwltlrp-----qelvx-----xgdlpislwlnhgyrclygshtag 106

101 WQXSKLEL 110

107 weepdpl 116

RESULT 29
 AAB04173
 ID AAB04173 standard; Protein: 289 AA.
 AC AAB04173;
 DT 11-APR-2001 (first entry)

Leukocyte immunoglobulin like receptor pmb36-2.

Leukocyte immunoglobulin like receptor; LIR; gene therapy;
 autoImmunity; autoimmune disorders; immune system; human; ds.

Homo sapiens.

MO20068383-A2.

16-NOV-2000.

12-MAY-2000: 2000MO-US13728.

12-MAY-1999: 99US-0310463.

(IMM) IMMUNEX CORP.

Cosman DJ, Anderson DM, Borges L;

WPI: 2000-607645/67.

N-PSDB: AA044597.

Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
 treating autoimmune diseases and disease states with suppressed immune
 function

Example 9: Page 79-80; 117pp; English.

Host cells transformed with a vector which are capable of
 expressing a leukocyte immunoglobulin like receptor (LIR), can be
 used to produce the LIR polypeptide. LIR coding sequences may be
 used in the gene therapy of disorders mediated directly or
 indirectly by defective or insufficient amounts of any of the LIR
 polypeptides. The LIR polypeptides can be used to treat autoimmune
 diseases and disease states with suppressed immune function.

Sequence 289 AA:

Query Match 11.2% Score 74; DB 21; Length 289;

Best Local Similarity 27.7% Pred. No. 1.4;

Matches 36; Conservative 15; Mismatches 45; Indels 34; Gaps 7;

1 MSMLVFLILGCVTGMPTGAIFETGXSIAMSESHXLTQ-----CDN----- 47

1 mprllvllcglstlprthvqagllprllwepgsavl-vgspvltvcgllletqeyr 59

48 ---DVPRPCDSRLPAVQENGAQEPYNIDSPAIKHOF---LITGTCGRYRCNSGLST-G 100

60 lyrekkcapwltlrp-----qelvx-----xgdlpislwlnhgyrclygshtag 106

101 WQXSKLEL 110

107 weepdpl 116

RESULT 30

AAB02544

20 AAB02544 standard; Protein: 650 AA.

AC AAB02544;

DT 09-FEB-1999 (first entry)

Human LIR-P362 protein.


```

Oy      61 VGE-----WGAGEVHLSDSAKIQF-----LLTDDIGQRCRSGSLST-CMX 102
Db      52 qetqeyrlytkektpvltlrpqcl-vksgqfipsltwelwagryrcyzsdags 108
          : | | | | |
Oy      103 QLSLLEL 110
          : | | | |
Db      109 esedplel 116

RESULT_32
ID       AA662782
AC       AA662782 standard; Protein: 651 AA.
XX
XX       AA662782:
XX
Dn      23-sep-1998 (first entry)
DE
XX
XX       Protein encoded by human monocyte gene designated KIM67.
XX
XX       Human: type I transmembrane protein; immunoglobulin-like domain;
XX       EPOB: activated monocyte; TEOJ, RT603; control; development;
XX       CD14: common leukine; immune system; treatment; cancerous condition;
XX       degenerative condition; inflammation; host disease; diagnosis;
XX       graft versus host disease; inflammatory condition; detection; diagnosis;
XX       drug screening.
XX
XX       Homo sapiens.
XX
XX       M09824906-A2.
XX
XX       11-JUN-1998.
PF
XX       05-DEC-1997:   97WO-US21101.
XX
PR       21-MAR-1997:   97US-0041279.
PR       06-DEC-1996:   96US-0032252.
PR       09-DEC-1996:   96US-0762187.
PR       16-DEC-1996:   96US-0033181.
PA
XX       (SCHRF ) SCHERING CORP.
PI
PI       Adams GT, Gorman DM, Lanier LJ, McClanahan TK, Meynard L;
PI       Phillips JH, Zurawski G, Zurawski SM;
DR
DR       WPI: 1998-333355/29.
DR
XX       N-P-SDB; AAV38608.
XX
XX       New isolated activated monocyte cell genes(s) - used to develop
XX       products for treating various degenerative conditions,
XX       autoimmune responses, transplant rejection or inflammatory
XX       conditions
XX
XX       Claim 5; Pages 98-100; 104pp; English.
XX
CC       AA662777-82 are encoded by huamn monocyte genes, collectively
CC       designated KIM03. The genes are found in activated monocytes. The
CC       specification also describes other proteins encoded by activation in
CC       peripheral blood mononuclear cells, and/or physiology of the
CC       controlling development, differentiation, and/or physiology of the
CC       mammalian immune system. The products can be used for treating abnormal
CC       proliferation, regeneration, degeneration or atrophy. They can be used
CC       for treating e.g., cancerous conditions, degenerative conditions,
CC       autoimmune responses, transplantation rejection, graft versus host
CC       disease, or inflammatory conditions. The products can also be used for
CC       detection, diagnosis and drug screening.
XX
XX       Sequence     651 AA:
XX
Query Match              11.2%; Score 74; DB 19; Length 651:
Beat Local Similarity    26.6%; Pred. No. 3.9:
Matches   34: Conservative 14; Mismatches      50: Indels   30: Gaps    5

```

```

Oy      1 NKSMLVFLLLMWSGCVTEVAIAFEYRQKSLMASESHKLXTLGGCCDAVDPCPGDSRLPA 60
Db      1 :::: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1: 1:
Dy      1 mpLpVLlclglsglsgrtshvqghilpkPLtaeagpsvl-----vgspylctcgg 51
Oy      61 VQE-----KKQAEPIHLSDSPAIKIQQF---LITGDQGRCRSGLST GKX 102
Db      52 gteqegryrlyrkckapwrltlpqel----vkkgfipalthebnhyrvcygaaatga 108
Oy          |   |||
Db      103 QLSKLEH 110
           109 easspiet 116

RESULT 33
ID AAM1238
AC AAM1238 standard; protein: 122 AA.
XX
CX AAM1238:
DX 09-JUN-1998 (first entry)
DE Alpha-v-beta-5 antagonistic peptide derived from chicken MMP-2.
DM Matrik metalloproteinase-2; MMP-2; alpha-v-beta-5 antagonist; treatment;
KM vitronectin receptor; inhibition; angiogenesis; tumour growth;
KM restenosis; neovascularisation.
KM
KM Synthetic.
OS Gallus sp.
XX MO9745447.AL.
XX 04-DEC-1997.
PD 30-MAY-1997: 97MO-US09039.
PE 31-MAY-1995: 96GS-0018723.
PR 31-MAY-1995: 96GS-0015869.
XX (SCRI ) SCDIPS RES INST.
XX Brooks P., Cheereah DA., Friedlander M;
PI WPI: 1998-041758/04.
DR Packaging material containing polypeptide antagonist of alphav,
PT betas integrin - used for inhibition of angiogenesis, and for
PT treating tumours, inflammation, eye diseases etc.
PX Claim 2: Page -: 117pp; English.
XX Peptides AAM1234-13 are derived from the chicken matrix
CC metalloproteinase-2 (MMP-2) protein (AAM1227). The present peptide is
CC derived from amino acids 516-637. The peptides are able to act as
CC inhibitors of alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a vitronectin receptor.
CC describes a novel labelled package that contains an inhibitor of
CC angiogenesis. It is an alpha-v-beta-5 antagonising polypeptide that binds
CC to vitronectin. This invention provides particularly useful means for the
CC treatment of MMP-2. The antagonists are used to inhibit angiogenesis in inflamed
CC tissue. In solid tumours or metastases, and in a wide range of ocular
CC disorders (e.g. diabetic or other forms of retinopathy, neovascular
CC glaucoma, or corneal transplants). They are particularly used to induce
CC regression or to inhibit growth of tumours. The alpha-v-beta-5
CC smooth muscle cells following angioplasty and to reduce blood supply to
CC vessels. This invention also provides particularly useful means for the
CC alpha epithelial growth factor or especially vascular endothelial growth
CC factor. note: this sequence does not appear in the specification. It was
CC created using information provided.
XX
```

Sequence 122 AA:

Query Match 11.0% Score 73; DB 19; Length 122;
Best Local Similarity 23.6%; Pred. No. 0.56; 37; Indels 34; Gaps
Matches 26; Conservative 13; Mismatches

OY 17 PYTEALIFVETQSLVAHSEIKLTKTGQADVDY-----PQCSTRLPVQWENKQAP 69
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 15 pDqekvtf-----agewywyaaandrgypkkltslglpDqvqlaaImwgn-- 65

OY 70 VHDSPAIKHQGLTLDGPDGRCHRSLSSTGMOKSKILLETKGVKLACS 119
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 66 -----xkyrlsgdygkvkm-----eakkmetadcpklids 97

RESULT 34
AA041090
ID AA041090 standard: Protein: 193 AA.
DE AM041090:
PF 08-JUN-1998 (first entry)
XX Chicken matrix metalloproteinase CHMP-2 (aa445-637).
DE Matrix metalloproteinase: MMP-2; CHMP-2; gelatinase; chicken;
KM angiogenesis; inhibitor; antitumor; integrin alpha-v beta-3;
KM fibroblast receptor; rheumatoid arthritis; tumour; metastasis;
KB diabetic retinopathy; macular degeneration; restenosis; therapy.
OS Gallus sp.
XX MW9745137-A1.
PN 04-DEC-1997.
XX 30-MAY-1997; 97MO-US09158.
PE 31-MAY-1996; 96GS-0018733.
PR 31-MAY-1996; 96US-0015869.
XX (SCRI) SCRAPPS RES INST.
PA Brooks P., Chereah DA.
PI WPI: 1998-032334/03.
PX MPI: 1998-032334/03.

Claim 2: Page 158-159; 234pp: English.

This polypeptide comprises amino acid residues 445-637 of chicken mature matrix metalloproteinase 2 (CHMP-2, see AA041111). It can be produced by recombinant methods such as PCR amplification (see AAT15502) of CHMP-2 coding sequence (see AA003995) and cloning into e.g. pEX-3X vector for expression in E. coli as a fusion protein or directly into angiotensin II mediated in vitro specific fibronectin receptor alpha-v beta-3, and that inhibition of alpha-v beta-3 function inhibits angiogenesis. Claimed antagonists of alpha-v beta-3 comprise C-terminal fragments (see AA041083-94) of human or chicken MM4-109(1-10), derivatised polypeptides, a monoclonal antibody or organic amine compound. The antagonists are used to inhibit angiogenesis in vivo and/or in vitro, e.g. in models of rheumatoid arthritis, solid tumours or metastases particularly to induce tumour regression or inhibit growth of tumours; and in ocular disorders such as diabetic retinopathy or macular degeneration (all claimed). They can also be used to treat restenosis caused by migration of smooth muscle cells following angioplasty and to reduce

CC	blood supply to selected tissues (claimed). The new antagonists are
CC	highly selective for angiogenesis. Only new blood vessels express
CC	alpha-v-beta-3, so mature vessels are unaffected, and the antagonists
CC	could be of low toxicity.
XX	
XX	Sequence 193 AA:
SQ	
Query Match	11.0% Score 73; DB 19; Length 193;
Best Local Similarity	23.6%; Pred. No. 1.1;
Matches 26; Conservative 13; Mismatches 37; Indels 34; Gaps 4;	
Dd	17 PFYDAIRFTQXSNAMSEHKLTAKTQQDADPG-----PQDSRLPAVDKADDP 69
Oy	70 VLDSPAIKIOFLTAGOGVGRCHSLSSTOMXSLKLELGGYLACG 119
Dd	60 pcdkavif-----gngwxywaandrgypkktslgipddrqladafmgrr- 110
Dd	111 -----kkytysgdvrykyn-----eeekhmeatprklads 142
RESULT 35	
ID	AAW41235 standard; protein: 193 AA.
AC	AAW41235:
DT	09-JUN-1998 (first entry)
DE	Alpha-v-beta-5 antagonistic peptide derived from chicken MMP-2.
KM	Matrix metalloproteinase-2: MMP-2; alpha-v-beta-5 antagonist; treatment;
KM	vitronectin receptor; inhibition; angiogenesis; tumour growth;
KM	restenosis; neovascularisation.
OS	Synthetic.
OS	GALUS sp.
PX	MO9745447-A1.
PD	04-DEC-1997.
PE	30-MAY-1997; J7WO-USO9009.
PR	31-MAY-1996; SGUS-0018273.
PR	31-MAY-1996; 96US-0015669.
PA	(SCRI) SCRIIPS RES INST.
F1	Brooks P, Chereah DA, Friedlander M;
X1	WPI: 1996-041758/04.
XX	
XX	
XX	
PS	Claim 2: Page - : 117pp: English.
XX	
XX	Peptides AAW41234-35 are derived from the chicken matrix
CC	metalloproteinase-2 (MMP-2) protein (AAM41127). The present peptide is
CC	derived from amino acids 445-637. The peptides are able to act as
CC	alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a vitronectin receptor.
CC	Inhibitors of alpha-v-beta-5 can inhibit angiogenesis. The specification
CC	describes a novel labelled package that contains an inhibitor of
CC	angiogenesis i.e. an alpha-v-beta-5 antagonising polypeptide that binds
CC	to integrin alpha-v-beta-5 and includes a part of the C-terminal domain
CC	of vitronectin. In solid tumours or metastases, and in a wide range of ocular
CC	diseases (e.g. diabetic or other forms of retinopathy, neovascular
CC	glaucoma, or corneal transplants). They are particularly used to induce
CC	regression or to inhibit growth of tumours. The alpha-v-beta-5
CC	antagonists can also be used to treat restenosis caused by migration of

CC smooth muscle cells following angioplasty and to reduce blood supply to
CC selected tissues. The antagonists particularly inhibit neovascularisation
CC where this is induced by cytokines, e.g. transforming growth factor- β
CC and is used in the treatment of atherosclerosis, diabetic retinopathy or
CC factor. note: this sequence does not appear in the specification. It was
CC created using information provided.

SO Sequence 193 AA:

Query Match 11.0% Score 73; DN 19; Length 193;
Best Local Similarity 23.6%; Pred. No. 1.1;

Matches 26; Conservative 13; Mismatches 37; Indels 34; Gaps 4;

OY 17 PVTPEAIPETQXSLMAESRHLTKTGCCADVPC-----PDCSRLEPANCQACQEP 69
| :
Db 86 pdqekavff-----gqneywytlsnldrypkltalsgldpdyqfdaefwngm-- 136
OY 70 VHLDSPAIKQIPLTGDQGRRCRSGSLXKXQSLKLEITGPVYLACS 119
| :
Db 137 -----kkytlfsgrdywkyk-----eeekkmelaipclfaads 168

RESULT 36

AAW41089
ID AAW41089 standard; Protein: 228 AA.

AAW41089:

08-JUN-1998 (first entry)

Chicken matrix metalloproteinase cMMP-2 (aa410-637).

Matrix metalloproteinase: MMP-2; cMMP-2; gelatinase; chicken;
KX collagenase; osteoblast growth inhibitor; osteoporosis;
KM vitronectin receptor; rheumatoid arthritis; tumour metastasis;
KN diabetic retinopathy; macular degeneration; restenosis; therapy.

Callus sp.

MO9745137-A1.

04-DEC-1997.

30-MAY-1997; 97MO-US09158.

31-MAY-1996; 96OUS-0018733.

31-MAY-1996; 96OUS-0015869.

(SCRI) SCRIPPS RES INST.

Brooks P, Chersosh DA;

WPI: 1998-032334/03.

PKaging material containing polypeptide antagonist of alphav,
PT beta3 integrin - used for inhibition of angiogenesis, and for
XX treating tumours, inflammation, eye diseases etc.
PS Claim 21: Page 157-158; 234pp; English.

This polypeptide comprises amino acid residues 410-637 of chicken
mature matrix metalloproteinase 2 (cMMP-2, see AAW41111). It can be
CC produced by recombinant methods such as PCR amplification (see
CC AAV12301) of cMMP-2 coding sequence (see AAV03995) and cloning into
CC a suitable expression vector, or by chemical synthesis of the peptide
CC with glutathione-S-transferase. The invention relates to the
CC discovery that angiogenesis is mediated by the specific vitronectin
CC receptor alpha-v beta-3, and that inhibition of alpha-v beta-3
CC function inhibits angiogenesis. Claimed antagonists of alpha-v
CC beta-3 comprise C-terminal fragments (see AAW41083-94) of human or
CC chicken MMP-2, fusion polypeptides, cyclic or linear polypeptides
CC (see also AAW41098-110), derivatised polypeptides, a monoclonal

CC antibody or organic mimetic compound. The antagonists are used to
CC inhibit angiogenesis in: inflamed tissue for treatment of arthritis
CC or rheumatoid arthritis; solid tumours or metastases, particularly
CC in the treatment of cancer; diabetic retinopathy or macular degeneration
CC or other disorders such as diabetic retinopathy or macular degeneration
CC (all claimed). They can also be used to treat restenosis caused by
CC migration of smooth muscle cells following angioplasty and to reduce
CC blood supply to selected tissues (claimed). The new antagonists are
CC highly selective for angiogenesis. Only new blood vessels express
CC alpha-v beta-3, so mature vessels are unaffected, and the antagonists
CC should be of low toxicity.

SO Sequence 228 AA:

Query Match 11.0% Score 73; DB 19; Length 228;
Best Local Similarity 23.6%; Pred. No. 1.3;

Matches 26; Conservative 13; Mismatches 37; Indels 34; Gaps 4;

OY 17 PVTPEAIPETQXSLMAESRHLTKTGCCADVPC-----PDCSRLEPANCQACQEP 69
| :
Db 95 pdqekavff-----gqneywytlsnldrypkltalsgldpdyqfdaefwngm-- 145
OY 70 VHLDSPAIKQIPLTGDQGRRCRSGSLXKXQSLKLEITGPVYLACS 119
| :
Db 146 -----kkytlfsgrdywkyk-----eeekkmelaipclfaads 177

RESULT 37

AAW41234
ID AAW41234 standard; Protein: 228 AA.

AAW41234:

09-JUN-1998 (first entry)

Alpha-v-beta-5 antagonistic peptide derived from chicken MMP-2.

Matrix metalloproteinase-2; MMP-2; alpha-v-beta-5 antagonist; treatment;
KX vitronectin receptor; inhibition; angiogenesis; tumour growth;
KM restenosis; neovascularisation.

Synthetic.

Callus sp.

MO9745447-A1.

04-DEC-1997.

30-MAY-1997; 97MO-US09099.

31-MAY-1996; 96OUS-0018733.

31-MAY-1996; 96OUS-0015869.

(SCRI) SCRIPPS RES INST.

Brooks P, Chersosh DA, Friedlander W;

WPI: 1998-041758/04.

PKaging material containing polypeptide antagonist of alphav,
PT beta3 integrin - used for inhibition of angiogenesis, and for
XX treating tumours, inflammation, eye diseases etc.
PS Claim 21: Page --, 117pp; English.

Peptides AAW41234-39 are derived from the chicken matrix
metalloproteinase-2 (MMP-2, protein (AAW41227)). The present peptide is
CC derived from amino acids 410-637. The peptides are able to act as
CC alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a vitronectin receptor.
CC Inhibitors of alpha-v-beta-5 can inhibit angiogenesis. The specification
CC describes a novel labelled package that contains an inhibitor of
CC angiogenesis i.e. an alpha-v-beta-5 antagonising polypeptide that binds

RESULT 42
 AAB04181
 ID AAB04181 standard; Protein: 287 AA.
 XX AAB04181:
 XX
 XX 11-APR-2001 (first entry)
 DE Leukocyte immunoglobulin like receptor LIR-9m2.
 XX
 XX Leukocyte immunoglobulin like receptor; LIR; gene therapy;
 XX autoimmunity; autoimmune disorders; immune system; human.
 XX Homo sapiens.
 XX
 XX W0200068383-A2.
 XX
 XX 16-NOV-2000.
 PD 12-MAY-2000: 2000MO-US13228.
 PP 12-MAY-1999: 99US-0310463.
 XX
 XX (IMM V) IMMUNEX CORP.
 PA
 PI Cosman DJ, Anderson DM, Borges L;
 XX
 XX WPI: 2000-687645/67.
 DR N-PSDB: AAB54609.
 XX
 XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
 PT treating autoimmune diseases and disease states with suppressed immune
 PT function
 XX
 XX Claim 1: Page 108-109; 11/Pp; English.
 XX
 XX Most cells transformed with a vector which are capable of
 CC expressing a leukocyte immunoglobulin like receptor (LIR) can be
 CC used to produce the LIR polypeptide. LIR coding sequences may be
 CC used in the gene therapy of disorders mediated directly or
 CC indirectly by defective or insufficient amounts of any of the LIR
 CC polypeptides. The LIR polypeptides can be used to treat autoimmune
 CC diseases and disease states with suppressed immune function.
 XX
 XX Sequence 287 AA:
 S0
 Query Match 11.0%; Score 72.5; DB 21; Length 287;
 Best Local Similarity 26.0%; Pred. No. 2;
 Matches 26; Conservative 13; Mismatches 34; Indels 27; Gaps 5;
 Oy 27 TOXSIAME-----SEHKLKTGGCCDAD-----VRCPPDSRLPAVQEMQAGEFVHLD 73
 Db 33 SKELTWEQGVSVISGNSVITITCGTCLAEQGVITVKEGSP-----EPKDCQNP1--- 81
 Oy 74 SPARKHOFLLTGOTO--GYYRCSCGLSTCWKQXLSKLEL 110
 Db 82 ephkxarfiptmhahgrfycyydpagvsepsdpl 121
 RESULT 43
 AABY1504
 ID AABY1504 standard; Protein: 542 AA.
 XX AABY1504:
 XX
 XX 25-JAN-2000 (first entry)
 XX
 XX Chimeric protein IRF-7(1-246)/IRF-3(132-427).
 XX
 XX Interferon regulatory factor; IRF; chimeric protein; serine; threonine;
 XX carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;
 XX post-translational modification; sendai virus; cancer treatment; herpes;
 KW

KW PRD1/PRD11; promoter; ISRE regulatory element; stimulation; activate;
 KW activator; DNA binding; transcriptional activity; viral infection;
 KW proteasome mediated degradation; Influenza; HIV infection; cytokine gene;
 KW target cell.
 XX Homo sapiens.
 XX
 XX Synthetic.
 XX
 XX Key
 XX Region
 FT Location/Qualifiers
 FT 1..246
 FT /note= "Corresponds to modified amino-terminal domain
 FT of IRF-7 protein"
 FT 247..542
 FT /note= "Corresponds to modified carboxy-terminus of
 FT IRF-3 protein"
 FT 178-3 protein
 FT Misc-difference 511
 FT /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 513
 FT /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 517
 FT /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 519
 FT /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 520
 FT /note= "Wild type Ser replaced with Asp"
 FT
 XX W09951737-A1.
 XX
 XX 14-OCT-1999.
 PD 07-APR-1999: 99MO-CAN0314.
 XX
 XX 07-APR-1998: 98CA-2234588.
 PR
 XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
 PA
 XX Hiscott J. Lin R;
 XX
 XX WPI: 1999-620201/53.
 DR N-PSDB: AAX90994.
 XX
 XX Carboxy-terminus modified highly active forms of interferon regulatory
 PT factor proteins used for the treatment of viral infections
 PT
 XX Claim 11: Fig 13; 9pp; English.
 XX
 XX The present sequence is the chimeric protein comprising, residues 1-246
 CC from the modified amino-terminal domain of interferon regulatory factor,
 CC IRF-7 and 132-427 residues from the carboxy-terminus of modified IRF-3
 CC (5D) protein. The serine and threonine residues are post-translationally
 CC modified by phosphorylation, following sendai virus infection. The
 CC modified IRF, substituted with aspartic acid, functions as a strong
 CC activator of promoters containing ISRE and PRD1/PRD11 regulatory
 CC elements. The modified protein is used for the treatment of viral and
 CC proteinase mediated degradation. The modified IRF sequences are used for
 CC the treatment of viral infections like, influenza, herpes or HIV
 CC infection. They may also be used to activate a cytokine gene, in cancer
 CC treatment or to modify a target cell of an organism.
 XX
 XX Sequence 542 AA:
 S0
 Query Match 10.9%; Score 72; DB 20; Length 542;
 Best Local Similarity 24.7%; Pred. No. 5.4;
 Matches 40; Conservative 14; Mismatches 56; Indels 53; Gaps 6;
 Oy 7 FULLANGVGTGTEALFETQSLVAESSEHKLKTGGCCDADVP---GPP----- 53
 Db 105 FWLIDISGQDPAQKHVYALSTELCEVGEQDGLQENPAPVPPGSGPAPGLAHDA 164
 Oy 54 -----GDSRLPAPVQ-----WGAQEPVHLSDPALKHQFL-ITGD 86
 Db 165 glapgpplpapagdkdilllqvvgscialhlltaswga-dpyvtekkapgeqgdlplvga 223


```

Best Local Similarity 25.5%; Pred. NO. 3.6;
Matches 28; Conservative 9; Mismatches 37; Indels 36; Gaps 5;
OY 4 LVPELLKGTGCPVTEA-----AIEETQXSLMESEHKLKTLGGCCADVP 51
    |||||
Db 12 lltcllwpygsewynkavltlqppwstlqgnvltlccgphl-----58
OY 52 PCGDSRIAPVQEW--GAQEPVHLDSFAIKHQFLTITDTCGRYRCRSGLS 98
    |||||
Db 59 -pyds--lqwlfnslavqlstpsyslpe--asfndsgseyrcjgss 100

```

Search completed: January 7, 2002, 16:49:15
 Job time: 307 sec

•
•
•
•


```

? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESS: Marshall, O'Toole, Gorstein, Murray & Borun
? STREET: 233 South Wacker Drive/6300 Sears Tower
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? ZIP: 60606
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: FASTED for Windows Version 2.0
? CURRENT APPLICATION DATA: base #1.0, Version #1.30
? FILING DATE: 20-FEB-1998
? APPLICATION NUMBER: US/09/402.002
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 1997-71082
? FILING DATE: 28-MAR-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: ECT/JP98/01286
? FILING DATE: 23-MAR-1998
? ATTORNEY/AGENT INFORMATION:
? NAME: Cavley, Jr., Thomas A.
? REGISTRATION NUMBER: 40,944
? REFERENCE/DOCKET NUMBER: 19036/36276
? TELECOMMUNICATION NUMBER:
? TELEPHONE: (312) 474-6300
? TELEFAX: (312) 474-0448
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1747 base pairs
? TYPE: nucleic acid
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Klebsiella pneumoniae
? STRAIN: Clinical isolate KP-85-43
? US-09-402-002-2

alignment_scores:
? Quality: 47.50 Length: 18
? Ratio: 3.167 Gaps: 1
? Percent Similarity: 83.333 Percent Identity: 44.444

Alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-402-002-2/rev ..
Align seg 1/1 to reverse of: US-09-402-002-2 from: 1 to: 1747
1 McSerMetLeuValaIlePheLeuLeu.....TFRGIYAlTrp 14
1442 CTCACAAATGATGATATTCCTGCTGCTGCATGACCTGGGACAC 1393
14 TTP 15
1392 ATGG 1389
seq_name: /cgn2_6/ptcdat2/1na/6A_COMB.seq:us-09-027-064-3
seq_documentation_block:
? Sequence 3, Application US/09027064
? Patent No. 6133016
? GENERAL INFORMATION:
? APPLICANT: BERGSMAN, USMAN
? APPLICANT: BERGSMAN, DEBK
? TITLE OF INVENTION: A YAK-1 RELATED SERINE/THREONINE
? TITLE OF INVENTION: PROTEIN KINASE-HTLA33
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:

```

```

? ADDRESS: RATNER & PRESTIA
? STREET: P.O. BOX 980
? CITY: VALLEY FORGE
? STATE: PENNSYLVANIA
? COUNTRY: USA
? ZIP: 19381
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTED for Windows Version 2.0
? CURRENT APPLICATION DATA: us-09-027-064
? FILING DATE: 20-FEB-1998
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/053,924
? FILING DATE: 28-JUL-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: PRESTIA, PAUL P.
? REGISTRATION NUMBER: 3,031
? REFERENCE/DOCKET NUMBER: GH-70172
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-407-0700
? TELEFAX: 610-407-0701
? TELEX: 946169
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1338 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-09-027-064-3

alignment_scores:
? Quality: 46.00 Length: 14
? Ratio: 4.182 Gaps: 0
? Percent Similarity: 78.571 Percent Identity: 57.143

Alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-027-064-3 ..
Align seg 1/1 to: US-09-027-064-3 from: 1 to: 1338
2 SerMetLeuValaIlePheLeuLeuTrpDLYVAlThTTP 15
845 TCCGCGCTCGTATGACCTGCGACCTGCTGCGGCGCTGACCTGG 886
seq_name: /cgn2_6/ptcdat2/1na/6B_COMB.seq:us-09-271-815-3
seq_documentation_block:
? Sequence 3, Application US/09271815
? Patent No. 6297036
? GENERAL INFORMATION:
? APPLICANT: BERGSMAN, DEBK
? APPLICANT: SHARON, USMAN
? TITLE OF INVENTION: A YAK-1 RELATED SERINE/THREONINE PROTEIN KINASE-HTLA33
? TITLE OF INVENTION: PROTEIN KINASE-HTLA33
? CURRENT FILING DATE: 1999-03-18
? EARLIER FILING DATE: 1998-02-20
? EARLIER FILING DATE: 1997-07-28
? NUMBER OF SEQ ID NOS: 6
? SPO ID NO: 3
? SPO ID NO: 3 Patent In Ver. 2.0
? LENGTH: 1338
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-271-815-3

```



```

alignment_scores:
  Quality: 46.00 Length: 14
  Ratio: 0.512 Gaps: 0
  Percent Similarity: 78.571 Percent Identity: 57.143

alignment_block:
  US-09-471-276-831_COPY_L16 x US-09-271-815-3
  Align seg 1/1 to: US-09-271-815-3 from: 1 to: 138

      2 sermetuevalvaipheuleuleutrcglyalvthtrp 15
      |||.....||| |||||.....|||
      845 TCCGCGCTCGTGTACCTGCGTCGACCTGGCGCTGG 886

seq_documentation_block:
  : Sequence 1, Application: US/09027064
  : Patent No. 6133006
  :
  : GENERAL INFORMATION:
  : APPLICANT: SHABON USAN
  : APPLICANT: BEGSGMA, DEK
  : TITLE OF INVENTION: A YAK-1 RELATED SERINE/THREONINE
  : TITLE OF INVENTION: PROTEIN KINASE-HTLRA33
  : NUMBER OF SEQUENCES: 6
  : COMPLETION DATE: 1998-02-26
  : ADDRESSEE: RAYNER & PRESTIA
  : STREET: P.O. BOX 980
  : CITY: VALLEY FORGE
  : STATE: PA
  : COUNTRY: USA
  : ZIP: 19482
  :
  : COMPUTER READABLE FORM:
  : MEDIUM TYPE: Diskette
  : OPERATING SYSTEM: Windows compatible
  : SOFTWARE: PASTED for Windows Version 2.0
  :
  : CURRENT APPLICATION DATA:
  : APPLICATION NUMBER: US/09/027,064
  : FILING DATE: 20-FEB-1998
  :
  : CLASSIFICATION:
  : PRIORITY APPLICATION DATA:
  : APPLICATION NUMBER: 60/053,924
  : FILING DATE: 28-JUL-1997
  : ATTORNEY REFERENCE: 100000000
  : NAME: /PRESTIA, PAUL F.
  : REGISTRATION NUMBER: 23,031
  : REFERENCE/DOC#ET NUMBER: GH-70172
  : TELECOMMUNICATION INFORMATION:
  : TELEPHONE: 610-407-0700
  : TELEFAX: 610-407-0701
  : TELEX: 846169
  : INFORMATION FOR SEQ ID NO: 1:
  : SOURCE: CHINA BATTERIES
  : SEQUENCE LENGTH: 2394 bases
  : TYPE: nucleic acid
  : STRANDEDNESS: single
  : TOPOLOGY: linear
  : MOLECULE TYPE: cDNA
  : US-09-027-064-1

alignment_scores:
  Quality: 46.00 Length: 14
  Ratio: 4.182 Gaps: 0
  Percent Similarity: 78.571 Percent Identity: 57.143

alignment_block:
  US-09-471-276-831_COPY_L16 x US-09-027-064-1
  Align seg 1/1 to: US-09-027-064-1 from: 1 to: 2394

```

```

seq_name: /cgn2/6/prodata/2/1na/6B_COMB_seq: US-09-271-815-1
seq_documentation_block:
  ? Sequence 1, Application US/09271 815
  ? Patent No. 6297036
  ? GENERAL INFORMATION:
  ? APPLICANT: BERKMAN, DEBK
  ? APPLICANT: SHAMON, USMAN
  ? TITLE OF INVENTION: A YAK-1 RELATED SERINE/THREONINE PROT
  ? FILE NUMBER: 6H-70112
  ? CURRENT FILING DATE: 1998-03/19
  ? EARLIER FILING DATE: 1998-03/19
  ? EARLIER APPLICATION NUMBER: 09/027,064
  ? EARLIER FILING DATE: 1998-02-20
  ? EARLIER APPLICATION NUMBER: 66/053,924
  ? EARLIER FILING DATE: 1997-07-28
  ? NUMBER OF SEQ ID NOS: 6
  ? SOFTWARE: PatentIn Ver. 2.0
  ? SEQ ID NO 1
  ? SEQ ID NO 2
  ? SEQ ID NO 3
  ? TYPE: DNA
  ? ORGANISM: Homo sapiens
  ? US-09-271-815-1

alignment_scores:
  ? Quality: 46.00 Length: 14
  ? Ref: 78.571 Gaps: 0
  ? Percent Similarity: 78.571 Percent Identity: 57.143

alignment:
  ? US-09-471-2716-831.COPY.1.16 x US-09-271-815-1 ..
  ? Align seg 1/1 to: US-09-271-815-1 from: 1 to: 2394
  ?
  ? 2 Set:lelelva1ValPheleulelelelelTfGCllyalTfTfTfP 15
  ? 1901 TCCGCCCTCTGATGACCTCCATCCATCTCGGGGCTCTACCTG 1942
  ?
  seq_name: /cgn2/6/prodata/2/1na/5B_COMB_seq: US-08-816-693A-1
seq_documentation_block:
  ? Sequence 1, Application US/0881693A
  ? Patent No. 5874241
  ? GENERAL INFORMATION:
  ? APPLICANT: Takahashi, Joseph S
  ? APPLICANT: Turck, Fred W
  ? APPLICANT: Plinto, Lawrence H
  ? TITLE OF INVENTION: Clock Gene and Gene Product
  ? NUMBER OF SEQUENCES: 53
  ? CORRESPONDENCE ADDRESS:
  ? ADDRESSEE: Dressler, Rocky, Milanow & Katz
  ? STREET: Two Prudential Plaza, Suite 1700
  ? CITY: CHICAGO
  ? STATE: ILLINOIS
  ? COUNTRY: USA
  ? ZIP: 60601
  ?
  ? COMPUTER READABLE FORM:
  ? MEDIUM TYPE: Floppy disk
  ? COMPUTER: IBM PC compatible
  ? OPERATING SYSTEM: PC-DOS/MS-DOS
  ? SOFTWARE: SeqMan, Release 11.0, Version #1.30
  ? CURRENT APPLICATION NUMBER: 08/016,934
  ? APPLICATION NUMBER: US/08/816,93A
  ? FILING DATE:
  ?
  ? CLASSIFICATION: 435
  ? ATTORNEY/AGENT INFORMATION: Thomas E
  ? NAME: No. 5674241.thrup, Thomas E

```

```

: REGISTRATION NUMBER: 33,268
: TELECOMMUNICATION INFORMATION:
: PATENT NO.: 6291429
: TELEPHONE: 312-615-6100
: TELEFAX: 312-615-6100
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7498 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MODIFICATION: none
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 389..2954
US-08-816-693A-1

alignment_scores:
Quality: 46.00 Length: 9
Ratio: 5.111 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-816-693A-1 ..
Align seq 1/1 to: US-08-816-693A-1 from: 1 to: 7498
8 LeuLeuLeuTrpGlyValThrTrpGly 16
|||||
5598 CTCCTCTCTATGCGGTGCTGCTGCGGT 5624

seq_name: /cgn2_6/p/cdata/2/lna/6A_COMB.seq:US-08-885-291-1

seq_documentation_block:
: Sequence 1, Application US/08885291A
: Patent No. 6057125
: GENERAL INFORMATION:
: APPLICANT: Takahashi, Joseph S.
: APPLICANT: Turck, Fred W.
: TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
: INVENTOR: Turck, Fred W.
: CURRENT APPLICATION NUMBER: US/08/885,291A
: CURRENT FILING DATE: 1997-06-30
: EARLIER APPLICATION NUMBER: 08/816,693
: EARLIER FILING DATE: 1997-03-13
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 7498
: TYPE: DNA
: ORGANISM: Mus musculus
US-08-885-291-1

alignment_scores:
Quality: 46.00 Length: 9
Ratio: 5.111 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-885-291-1 ..
Align seq 1/1 to: US-08-885-291-1 from: 1 to: 7498
8 LeuLeuLeuTrpGlyValThrTrpGly 16
|||||
5598 CTCCTCTCTATGCGGTGCTGCTGCGGT 5624

seq_name: /cgn2_6/p/cdata/2/lna/6B_COMB.seq:US-09-456-672-1

seq_documentation_block:
: Sequence 1, Application US/09456672
: Patent No. 6291429
: TELECOMMUNICATION INFORMATION:
: PATENT NO.: 6291429
: TELEPHONE: 312-615-6100
: TELEFAX: 312-615-6100
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7498 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MODIFICATION: none
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 389..2954
US-09-456-672-1

alignment_scores:
Quality: 46.00 Length: 9
Ratio: 5.111 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-456-672-1 ..
Align seq 1/1 to: US-09-456-672-1 from: 1 to: 7498
8 LeuLeuLeuTrpGlyValThrTrpGly 16
|||||
5598 CTCCTCTCTATGCGGTGCTGCTGCGGT 5624

seq_name: /cgn2_6/p/cdata/2/lna/6B_COMB.seq:US-09-172-108-48

seq_documentation_block:
: Sequence 48, Application US/09172108
: Patent No. 6160104
: GENERAL INFORMATION:
: APPLICANT: Ziegler, Gary B.
: APPLICANT: Panzer, Scott R.
: APPLICANT: Selheimer, Jeffrey J.
: TITLE OF INVENTION: MARKERS FOR PEROXISOMAL PROLIFERATORS
: INVENTOR: Ziegler, Gary B.
: CURRENT APPLICATION NUMBER: US/09/172,108
: CURRENT FILING DATE: 1996-10-13
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PERL Program
: SEQ ID NO 48
: LENGTH: 285
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 706607235H1
US-09-172-108-48

alignment_scores:
Quality: 45.00 Length: 16
Ratio: 3.750 Gaps: 0
Percent Similarity: 75.000 Percent Identity: 56.250

alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-172-108-48 ..
Align seq 1/1 to: US-09-172-108-48 from: 1 to: 285
1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGly 16
|||||
|||||
```

```

: PATENT NO. 6291429
: GENERAL INFORMATION:
: APPLICANT: Takahashi, Joseph S.
: APPLICANT: Turck, Fred W.
: APPLICANT: Puro, Lawrence H.
: TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
: INVENTOR: Turck, Fred W.
: CURRENT APPLICATION NUMBER: US/09/496,672
: CURRENT FILING DATE: 2000-02-03
: PRIOR APPLICATION NUMBER: 08/885,291
: PRIOR FILING DATE: 1997-06-30
: PRIOR APPLICATION NUMBER: 08/816,693
: PRIOR FILING DATE: 1997-03-13
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 7498
: TYPE: DNA
: ORGANISM: Mus musculus
US-09-456-672-1

alignment_scores:
Quality: 46.00 Length: 9
Ratio: 5.111 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 88.889

alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-456-672-1 ..
Align seq 1/1 to: US-09-456-672-1 from: 1 to: 7498
8 LeuLeuLeuTrpGlyValThrTrpGly 16
|||||
5598 CTCCTCTCTATGCGGTGCTGCTGCGGT 5624

seq_name: /cgn2_6/p/cdata/2/lna/6B_COMB.seq:US-09-172-108-48

seq_documentation_block:
: Sequence 48, Application US/09172108
: Patent No. 6160104
: GENERAL INFORMATION:
: APPLICANT: Ziegler, Gary B.
: APPLICANT: Panzer, Scott R.
: APPLICANT: Selheimer, Jeffrey J.
: TITLE OF INVENTION: MARKERS FOR PEROXISOMAL PROLIFERATORS
: INVENTOR: Ziegler, Gary B.
: CURRENT APPLICATION NUMBER: US/09/172,108
: CURRENT FILING DATE: 1996-10-13
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PERL Program
: SEQ ID NO 48
: LENGTH: 285
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 706607235H1
US-09-172-108-48

alignment_scores:
Quality: 45.00 Length: 16
Ratio: 3.750 Gaps: 0
Percent Similarity: 75.000 Percent Identity: 56.250

alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-172-108-48 ..
Align seq 1/1 to: US-09-172-108-48 from: 1 to: 285
1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGly 16
|||||
|||||
```

```
14 ATGCTCTCTGACACTGACACTGACTCTCTCGGGTTTCATCTCGGAC 61
seq_name: /cgn2_6/p/odata/2/lna/5B.COMB.seq:us-09-172-711-45
seq_documentation_block:
: Sequence 1 Application US/09172711
: GENERAL INFORMATION:
: PATENT NO. 6160105
: APPLICANT: Cunniffham, Mary Jane
: APPLICANT: Zweiger, Gary B.
: APPLICANT: Panzer, Scott R.
: APPLICANT: Seilheimer, Jeffrey J.
: TITLE OF INVENTION: IDENTIFYING TOXICOLOGICAL RESPONSES
: FILE REFERENCE NO.: 001115
: CURRENT FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: PERL Program
: SEQ ID NO 45
: LENGTH: 285
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: 70607235H1
US-09-172-711-45

alignment_scores:
: Quality: 45.00 Length: 16
: Ratio: 3.750 Gaps: 0
: Percent Similarity: 75.000 Percent Identity: 56.250

alignment_block:
US-09-471-276-831_copy_1_16 x US-09-172-711-45 ..
Align seq 1/1 to: US-09-172-711-45 from: 1 to: 285
1 MetSerMetLeuValAlaPheLeuLeuLeuTrpGlyValIleTrpGly 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14 ATGCTCTCTGACACTGACACTGACTCTCTCGGGCTTCGACATTCGAGC 61

seq_name: /cgn2_6/p/odata/2/lna/5A.COMB.seq:us-08-254-573-1
seq_documentation_block:
: Sequence 1 Application US/08254573
: GENERAL INFORMATION:
: PATENT NO. 5610032
: APPLICANT: KAMBOJ, Rajender
: APPLICANT: ELIOTT, Candace
: APPLICANT: NUTT, Stephen
: TITLE OF INVENTION: AMPA-BINDING HUMAN GLUR1 RECEPTORS
: FILE REFERENCE NO.: 001115
: CURRENT FILING DATE: 1999-03-08
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patenceln ver. 2.0
: SEQ ID NO 3
: LENGTH: 840
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(840)
US-08-254-419C-3

alignment_scores:
: Quality: 44.00 Length: 13
: Ratio: 4.400 Gaps: 0
: Percent Similarity: 76.923 Percent Identity: 53.846

alignment_block:
US-09-471-276-831_copy_1_16 x US-09-264-419C-3 ..
Align seq 1/1 to: US-09-264-419C-3 from: 1 to: 840
```

```
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16777/179 ALLE
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: E-MAIL: 999@alle.com
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3220 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 62..2782
: NAME/KEY: sig_peptide
: LOCATION: 62..115
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 116..2782
US-08-254-573-1

alignment_scores:
: Quality: 45.00 Length: 16
: Ratio: 3.462 Gaps: 0
: Percent Similarity: 81.250 Percent Identity: 50.000

alignment_block:
US-09-471-276-831_copy_1_16 x US-08-254-573-1/rev ..
Align seq 1/1 to reverse of: US-08-254-573-1 from: 1 to: 3220
1 MetSerMetLeuValAlaPheLeuLeuLeuTrpGlyValIleTrpGly 16
2875 CTTCATTTCTGTTGTTGTTGTTGTCACCTGACGAGGTTGCGATCGAGC 2828
```

```
seq_name: /cgn2_6/p/odata/2/lna/5B.COMB.seq:us-09-264-419C-3
seq_documentation_block:
: Sequence 3, Application US/09264419C
: GENERAL INFORMATION:
: PATENT NO. 6174682
: APPLICANT: Rhododoust, Mehran
: TITLE OF INVENTION: No. 6174682a1 Thioresdoxin Family Active Site Molecules and us
: FILE REFERENCE: MN1-076
: CURRENT APPLICATION NUMBER: US/09/264,419C
: CURRENT FILING DATE: 1999-03-08
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patenceln ver. 2.0
: SEQ ID NO 3
: LENGTH: 840
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(840)
US-09-264-419C-3

alignment_scores:
: Quality: 44.00 Length: 13
: Ratio: 4.400 Gaps: 0
: Percent Similarity: 76.923 Percent Identity: 53.846

alignment_block:
US-09-471-276-831_copy_1_16 x US-09-264-419C-3 ..
Align seq 1/1 to: US-09-264-419C-3 from: 1 to: 840
```


alignment_scores:

Quality: 43.00 Length: 11
Ratio: 4.778 Gaps: 0
Percent Similarity: 81.618 Percent Identity: 54.545

alignment_block:

US-09-471-276-831_COPY_1_16 x US-08-931-608A-1/rev ..

Align seg 1/1 to reverse of: US-08-931-608A-1 from: 1 to: 3487

6 ValphaLeuValphaLeuTgPolValphaLeuTgPol 16

2757 ATTGCAATTTCATCTGGAAGATGATTCGGCA 2725

seq_name: /cgn2_6/pdata/2/lna/6A.COMB.seq:US-09-033-428-2

seq_documentation_block:

Sequence 2, Application US/05033428
Patent No. 6254862
GENERAL INFORMATION:
APPLICANT: Lamparski, Andrew
APPLICANT: Maltile, Henry
APPLICANT: Schuur, Eric
APPLICANT: Henderson, Daniel
TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC FOR CELLS
TITLE OF INVENTION: EXPRESSING APHA-PESTOPROTEIN AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 23
CLASSIFICATION: 435.970N
ADDRESSER: MORRISON & FORSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,428
FILING DATE: 05/09/93
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: POLITZI, CATHERINE M.
REGISTRATION NUMBER: 40,130
REFERENCE/DOCKET NUMBER: 34802-30004, 00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
FILER: 06141 MESSNICKS SFO
INFORMATION FOR SEQUENCE 1:
SEQUENCE CHARACTERISTICS: 2:
LENGTH: 5224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-428-2

alignment_scores:

Quality: 43.00 Length: 15
Ratio: 3.071 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 46.667

alignment_block:

US-09-471-276-831_COPY_1_16 x US-09-033-428-2/rev ..

Align seg 1/1 to reverse of: US-09-033-428-2 from: 1 to: 5224

2 SermetLeuValphaLeuLeuTgPolValphaLeuTgPol 16

.....

5139 AACCTTGTGCTATACCTGTAATGATTCCTTACCTAGCT 5095

seq_documentation_block:

Sequence 1, Application US/08676169
Patent No. 5773235
GENERAL INFORMATION:
APPLICANT: Chirnside, Ewan Douglas
TITLE OF INVENTION: EQUINE ARTERITIS VIRUS PEPTIDES; ANTIBODIES
TITLE OF INVENTION: EQUINE ARTERITIS VIRUS PEPTIDES; ANTIBODIES
NUMBER OF SEQUENCES: 7
CLASSIFICATION: 7
ADDRESSER: WATSON & VANDERHIE P.C.
STREET: 1100 NO. 5773235th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A
ZIP: 22201-4714
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,169
FILING DATE: 31-JUL-96
CLASSIFICATION: 435.970N
ADDRESSER: CREWFOORD, ARTHUR R.
NAME: CREWFOORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1498-85
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQUENCE 1:
SEQUENCE CHARACTERISTICS: 1:
LENGTH: 12687 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-676-169-1

alignment_scores:

Quality: 43.00 Length: 12
Ratio: 3.583 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 58.333

alignment_block:

US-09-471-276-831_COPY_1_16 x US-08-676-169-1 ..

Align seg 1/1 to: US-08-676-169-1 from: 1 to: 12687

1 MesSermetLeuValphaLeuLeuTgPolValphaLeuTgPol 12

.....

11132 TTATCTATGAGATTCATCTGATTCCTTCCTTGGCGCT 11167

seq_documentation_block:

Sequence 1, Application US/08981459
Patent No. 6090390
GENERAL INFORMATION:
APPLICANT: CHIRNSIDE, Ewan Douglas
TITLE OF INVENTION: Diagnostic Test For Equine Arteritis Virus
TITLE OF INVENTION: Diagnostic Test For Equine Arteritis Virus
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: 400 Seventh Street, N.W.
CITY: Washington

```

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/981,459
FILING DATE: 19-FEB-1998
CLASSIFICATION INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: HOLMAN, John C.
REGISTRATION NUMBER: 22,769
REFERENCE/DOCKET NUMBER: P61784050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-658-6666
FAX: 202-353-5350
TELEFAX: 202-353-5350
SEQUENCE IDENTIFICATION:
SEQUENCE CHARACTERISTICS: 1:
LENGTH: 12687 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-981,459-1

```

```

Alignment_scores:
Quality: 43.00      Length: 12
Ratio: 3.583      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 58.333

```

Alignment block:

US-09-471-276-831_COPY_1_16 x US-08-981-459-1 ...

Align seg 1/1 to: US-08-981-459-1 from: 1 to: 12687

1 MetcSerGctGauValpHeLcGauLentTpgLy 12

11133 TTACGTCAGTGGTGTACGTCATCTCTCTTGCGGT 11167

seq_name: /cgn2_6/prodata/2/lns/66_COMB.seq:US-09-328-111-517

seq_documentation_block:

Sequence 517, Application US/09328111

Patent No. 6262333

CLASSIFICATION:

APPLICANT: Steinmann, Allison O.

APPLICANT: Aetle, Jon H.

APPLICANT: Steinmann, Kathleen E.

APPLICANT: Burgess, Christopher C.

APPLICANT: Bushnell, Steven E.

APPLICANT: Carroll III, Eddie

APPLICANT: Celino, Theodore J.

APPLICANT: Dettl, Adam

APPLICANT: Ford, David M.

APPLICANT: Loxley, Maricle E.

APPLICANT: Monahan, John E.

APPLICANT: Schlegel, Robert

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

TITLE OF INVENTION: PRODUCTS

FILE REFERENCE: CCD-257 (US)

CURRENT APPLICATION NUMBER: US/09/328,111

CURRENT APPLICATION NUMBER: US/09/328,111

EARLIER FILING DATE: 1998-06-10

NUMBER OF SEQ ID NOS: 850

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 517

LENGTH: 611

TYPE: DNA

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(611)
OTHER INFORMATION: n..A..T..C.. or G
US-09-328-111-517

```

Alignment_scores:

```

Quality: 42.50      Length: 15
Ratio: 3.542      Gaps: 1
Percent Similarity: 80.000      Percent Identity: 66.667

```

Alignment block:

US-09-471-276-831_COPY_1_16 x US-09-328-111-517 ...

Align seg 1/1 to: US-09-328-111-517 from: 1 to: 611

3 MetcGauValpHeLcGauLentTpgLy 16

361 ATGCGTCAGTGGTGTACGTCATCTCTCTTGCGGT 405

seq_name: /cgn2_6/prodata/2/lns/5A_COMB.seq:US-08-583-318-4

seq_documentation_block:

Sequence 4, Application US/08583318

Patent No. 6563483

CLASSIFICATION:

APPLICANT: Staunton, Donald

APPLICANT: Harris, Edith

TITLE OF INVENTION: Cytoplasmic Modulators of Integrin

TITLE OF INVENTION: Binding

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSES: Marshfield, O'Toole, Gerstlein, Murray & Borun

Street: 333 South Wacker Drive, 6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

Medium type: floppy disk

Operating system: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/583,318

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/3050

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1437

US-08-583-318-4

```

Alignment_scores:
Quality: 42.50      Length: 15
Ratio: 3.542      Gaps: 1
Percent Similarity: 80.000      Percent Identity: 66.667

```



```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 18-NOV-1997
CLASSIFICATION NUMBER: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,040
FILING DATE: 18-NOV-1996
PRIOR APPLICATION DATA: US 60/061,328
APPLICATION NUMBER: 08-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Doyle Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-1202
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
FAX: 215-967-2591
TELEX: 831-494
SEQUENCE CHARACTERISTICS:
SEQUENCE FOR SEQ ID NO: 2:
LENGTH: 9936 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
NO. OF SEQUENCES: 1
NO. OF TYPES: DNA (genomic)
US-08-972-927-2

alignment_scores:
Quality: 42.50 Length: 15
Ratio: 3.542 Gaps: 1
Percent Similarity: 80.000 Percent Identity: 53.333

alignment_block:
US-09-471-276-831_Copy_1_16 x US-08-972-927-2 ..
Align seg 1/1 to: US-08-972-927-2 From: 1 to: 9936
1 MetSettelleValIvalPheLeuLeuUrrpGlyValThrTP 15
951 ATGTCAAGTTCGTTCTCATATACACTGTGTGGGGATA...TGG 992
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
seq_documentation_block:
Sequence 5, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
INVENTOR: Giesler, Karl J.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guesler, Karl J.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF STRUNDS: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP CODE: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

```



```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/276,531
? FILING DATE: HEREWITH
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? PRIORITY NUMBER: 50/079,677
? FILING DATE: March 27, 1998
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Lynn E. Murty, Ph.D.
? REGISTRATION NUMBER: PA-0008 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 845-4166
? TELEFAX: (650) 845-4166
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1350 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? IMMEDIATE SOURCE:
? LIBRARY: PANCYTUT01
? CLONE: 1514169
? US-09-276-531-5

alignment_scores:
? Quality: 42.00 Length: 9
? Ratio: 5.250 Gaps: 0
? Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:
? US-09-471-276-831_COPY_1_16 X US-09-276-531-5 ..
?
? Align seq 1/1 to: US-09-276-531-5 from: 1 to: 1350
?
?      8 leu1e4uutrrg1yvaltrtpg1y 16
?      |||||:::||||:::|||||||
?      726 TTCGTGATGCTGACTCATCATGCGGG 752

seq_name: /cgn2_6/p/rotate/2/1na/5a.COMB.seq:US-08-632-470-53
seq_documentation_block:
? Sequence 53 Application US/08632470
? Patent No. 5976791
? GENERAL INFORMATION:
? APPLICANT: MABILAT, CLAUDE
? APPLICANT: RAOUULT, DIDIER
? TITLE OF INVENTION: NUCLEOTIDE FRAGMENTS CAPABLE OF
? HYBRIDIZING SPECIFICALLY TO RICKETTSIA RNA OR RRNA AND
? NUMBER OF SEQUENCES: 1
? CURRENT APPLICATION DATA:
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: OLIF & BERRIDGE
? STREET: P.O. BOX 19928
? CITY: ALEXANDRIA
? STATE: VA
? COUNTRY: USA
? TELEPHONE: 703-836-6400
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/632,470
? PRIORITY NUMBER: 50/079,677
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: BERRIDGE, WILLIAM P
? REGISTRATION NUMBER: 30,024
? REFERENCE/DOCKET NUMBER: WPD 38238

```

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)-836-6400
? TELEFAX: (703)-836-2787
? INFORMATION FOR SEQ ID NO: 53:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1485 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-632-470-53

alignment_scores:
? Quality: 42.00 Length: 6
? Ratio: 7.000 Gaps: 0
? Percent Similarity: 100.000 Percent Identity: 83.333

alignment_block:
? US-09-471-276-831_COPY_1_16 X US-08-632-470-53 ..
?
? Align seq 1/1 to: US-08-632-470-53 from: 1 to: 1484
?
?      11 trfclYwalthrtpg1y 16
?      |||||:::|||||||
?      143 TGGCGATACCTCGCGA 160

seq_name: /cgn2_6/p/rotate/2/1na/5a.COMB.seq:US-08-299-810A-27
seq_documentation_block:
? Sequence 27 Application US/0829810A
? Patent No. 5721097
? GENERAL INFORMATION:
? APPLICANT: Rossau, Rudl
? APPLICANT: Van Heuverswyn, Hugo
? TITLE OF INVENTION: HYBRIDIZATION PROBES FOR THE
? NUMBER OF SEQUENCES: 26
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Merchant & Gould
? STREET: 3100 No. 5721097west Center
? CITY: Minneapolis
? STATE: MN
? COUNTRY: USA
? TELEPHONE: 554-0400
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/299,810A
? PRIORITY NUMBER: 08/1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hillson, Randall A.
? REGISTRATION NUMBER: 31,838
? REFERENCE/DOCKET NUMBER: 8076, 70-US-WO
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 612-332-5300
? TELEFAX: 612-332-5051
? INFORMATION FOR SEQ ID NO: 27:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1485 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORGANISM: Brachella catarrhalis
? IMMEDIATE SOURCE:
? CLONE: 16S rRNA Gene
? US-08-299-810A-27

```

alignment_scores:
 Quality: 42.00 Length: 6
 Ratio: 7.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 83.333

Alignment_block:

US-09-471-276-831_COPY_1_16 x US-08-299-810A-27 ..

Align seg 1/1 to: US-08-299-810A-27 from: 1 to: 1485

11 TCGGCAATCTGGGGA
 |||||:::|||||||
 95 TCGGCAATCTGGGGA 112

seq_name: /cgn2_6/prodata/2/lna/5b.COMB.seq:US-08-864-224-1

seq_documentation_block:

1. Application US/08864224
 Patent No. 5851808
 GENERAL INFORMATION:
 APPLICANT: Ellledge, Stephen J.
 APPLICANT: Liu, Qinghua
 TITLE OF INVENTION: Rapid Subcloning Using Site-Specific
 TITLE OF INVENTION: Recombination
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Medion & Carroll, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/864,224
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Ingolia, Diane E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: BCM-02681
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 307-8338
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS: 1:
 LENGTH: 2220 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: other nucleic acid
 RECOMBINATION: /desc .. DNA
 US-08-864-224-1

alignment_scores:

Quality: 42.00 Length: 9
 Ratio: 5.250 Gaps: 0
 Percent Similarity: 88.889 Percent Identity: 66.667

Alignment_block:

US-09-471-276-831_COPY_1_16 x US-08-864-224-1 ..

Align seg 1/1 to: US-08-864-224-1 from: 1 to: 2220

8 LeuLeuLeuTPrGlyValTTrPrGly 16
 ::|||:::||||||| |||||

673 ATCTATCTCTGGGGCTGGCTGGCCG 699

seq_name: /cgn2_6/prodata/2/lna/6a.COMB.seq:US-08-801-092-5

seq_documentation_block:

1. Sequence 5, Application US/08801092
 Patent No. 6074850
 GENERAL INFORMATION:
 APPLICANT: Antelman, Douglas
 APPLICANT: Gregory, Richard J.
 APPLICANT: Wils, Kenneth N.
 TITLE OF INVENTION: Tissue Specific Expression of
 TITLE OF INVENTION: Ribonucleoside Protein
 NUMBER OF SEQUENCES: 46
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Townsend and Townsend and CREW LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/801,092
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/751,517
 FILING DATE: 15-NOV-1996
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: US Inverness A, 35,136
 REGISTRATION NUMBER: 016930-001020
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 703-576-0300
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3583 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 209..250
 NAME/KEY: CDS
 LOCATION: 254..289
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 293..505
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 509..514
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 518..520
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 524..658
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 662..691
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 695..748

```

? NAME/KEY: CDS
? LOCATION: 752..781
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 785..829
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1132..1134
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1138..1149
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 833..862
? US-08-801-092-5

alignment_scores:
? Quality: 42.00 Length: 9
? Ratio: 5.250 Gaps: 0
? Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-801-092-5 ..
Align seg 1/1 to: US-08-801-092-5 from: 1 to: 3853
8 LeuLeuLeuTrpGlyValThrTrpGly 16
1307 ATTCTATTCGGGGGGGTGGGGTGGGCG 1333

seq_name: /cgn2_6/prodata/2/lna/6A.COMB.seq:US-08-801-092-19
seq_documentation_block:
? Sequence 19, Application US/08801092
? Patent No. 6074850
? GENERAL INFORMATION:
? APPLICANT: Artelman, Douglas
? APPLICANT: Gregory, Richard J.
? TITLE OF INVENTION: Tissue Specific Expression of
? NUMBER OF SEQUENCES: 46
? CORRESPONDENCE ADDRESS:
? ADDRESS: TOWNSEND AND TOWNSEND AND CREW LLP
? STREET: Two Embarcadero Center, 8th Floor
? CITY: San Francisco
? STATE: CA
? COUNTRY: USA
? ZIP: 94111

COMPUTER READABLE FORM:
? Sequence 19, Application US/08801092
? Patent No. 6074850
? COMPUTER-IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/801.092
? FILING DATE: 14-FEB-1997
? CLASSIFICATION: 514
? PRIORITY DATE: 15-NOV-1996
? APPLICATION NUMBER: US 08/751,517
? FILING DATE: 15-NOV-1996
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Fitts, Renee A.
? REGISTRATION NUMBER: 35,136
? REFERENCE/POCKET NUMBER: 0169930-001020
? TELEPHONE: 415-576-0300
? TELEFAX: 703-576-0300
? INFORMATION FOR SEQ ID NO: 19:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4026 base pairs

```

```

? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 209..250
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 254..289
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 293..505
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 509..514
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 518..520
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 524..658
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 662..691
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 695..748
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 752..781
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 785..829
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 833..862
? NAME/KEY: CDS
? LOCATION: 1305..1307
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1311..1322
? US-08-801-092-19

alignment_scores:
? Quality: 42.00 Length: 9
? Ratio: 5.250 Gaps: 0
? Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-801-092-19 ..
Align seg 1/1 to: US-08-801-092-19 from: 1 to: 4026
8 LeuLeuLeuTrpGlyValThrTrpGly 16
1480 ATTCTATTCGGGGGGGTGGGGTGGGCG 1506

seq_name: /cgn2_6/prodata/2/lna/6A.COMB.seq:US-08-801-092-33
seq_documentation_block:
? Sequence 33, Application US/08801092
? Patent No. 6074850
? GENERAL INFORMATION:
? APPLICANT: Artelman, Douglas
? APPLICANT: Gregory, Richard J.
? TITLE OF INVENTION: Tissue Specific Expression of
? NUMBER OF SEQUENCES: 46
? CORRESPONDENCE ADDRESS:

```

ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: 08/801.092
 PENDING APPLICATION NUMBER: 08/801.092
 CLASSIFICATION: 514
 FILING DATE: 15-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/751,517
 FILING DATE: 15-NOV-1996
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME/KEY: 254..289
 REGISTRATION NUMBER: A, 35,136
 REFERENCE/DOCKET NUMBER: 016930-001020
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-576-0200
 TELEFAX: 703-576-0300
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4283 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE: CDS
 LOCATION: 209..250
 NAME/KEY: CDS
 LOCATION: 254..289
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 293..505
 NAME/KEY: CDS
 LOCATION: 509..514
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 518..520
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 524..658
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 662..691
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 693..748
 NAME/KEY: CDS
 LOCATION: 753..761
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 785..829
 NAME/KEY: CDS
 LOCATION: 831..862
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1528..1530
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1534..1545
 US-08-801-092-33

alignment_scores:
 Quality: 42.00 Length: 9
 Ratio: 5.250 Gaps: 0
 Percent Similarity: 88.869 Percent Identity: 66.667
 alignment_block:
 US-09-471-276-831_copy_1_16 x US-08-801-092-33 ..
 Alignn seq 1/1 to: US-08-801-092-33 from: 1 to: 4249
 8 LeuLeuLeuTTCGCTGATTTTCGCTG 16
 1703 ATTCTATTCCTGGAGCTGGCTGGAGC 1729
 seq_name: /cgn2/6/plodata2/lna/5X_COWB.seq:US-08-343-401A-3
 seq_documentation_block:
 : Sequence 3, Application US/08343401A
 : Patent No. 5661132
 : CURRENT APPLICATION:
 : APPLICANT: Sealy, William F
 : APPLICANT: Macklin, Michael D
 : APPLICANT: Eriksson, Elof
 : APPLICANT: Andree, Christophe
 : TITLE OF INVENTION: Improved Wound Healing
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSER: Dr. Brady
 : STREET: PO Box 2113
 : CITY: Madison
 : STATE: WI
 : COUNTRY: USA
 : ZIP: 53701-2113
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : OPERATING SYSTEM: IBM PC compatible
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/343,401A
 : FILING DATE: 22-NOV-1994
 : CLASSIFICATION: 514
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Sealy, Nicholas J
 : REGISTRATION NUMBER: 27,386
 : REFERENCE/DOCKET NUMBER: 11-229-9103-9
 : TELEPHONE: 608-221-5000
 : TELEFAX: 608-221-5166
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4283 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: circular
 : MOLECULE TYPE: DNA (genomic)
 : INVARIANT SPACE
 : CLONE: pMHC1630
 : FEATURE:
 : NAME/KEY: exon
 : LOCATION: 713..721
 : FEATURE:
 : NAME/KEY: exon
 : LOCATION: 981..1253
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: join(713..721, 981..1253)
 : FEATURE:
 : NAME/KEY: s19_peptide
 : LOCATION: 713..1049
 : US-08-343-401A-3

```

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250      Gaps: 0
  Percent Similarity: 88.889      Percent Identity: 66.667

alignment_block:
  US-09-471-276-831_COPY_1_16 x US-08-343-401A-3  ..
Align seg 1/1  to: US-08-343-401A-3  from: 1  to: 4283
      8 leu1e0u1e0t7p1c3yvaltrh1t7p1c3y 16
      ::::::::::::::::::::::::::::::::::::
      1480 ATTCTATTCTGCGGGGTGCGGTGCGGCGC 1506

seq_name: /cgn2_6/p1c0data/2/1na/5A.COMB.seq:US-08-445-265A-1
seq_documentation_block:
  Sequence 1, Application US/08445265A
  Patient No. 5697901
  GENERAL INFORMATION:
  APPLICANT: Eriksson, Elof
  TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
  NUMBER OF SEQUENCES: 4
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Charles A Brady
  STREET: 1 South Pritchney Street
  CITY: Madison
  STATE: WI
  COUNTRY: US
  ZIP: 53703
  COMPUTER READABLE FORM:
  MEDIUM TYPE: floppy disk
  OPERATING SYSTEM: IBM PC COMPATIBLE
  SOFTWARE: PC-DOS/MS-DOS
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/445.265A
  FILING DATE:
  CLASSIFICATION: 604
  ATTORNEY/AGENT INFORMATION:
  NAME: Seay, Nicholas
  REGISTRATION NUMBER: 37186
  REFERENCE/DOCKET NUMBER: 110229.91080
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 608-251-5000
  TELEFAX: 608-251-9166
  INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
  TYPE: 283 base pairs
  STRANDEDNESS: double
  TOPOLOGY: circular
  MOLECULE TYPE: other nucleic acid
  FEATURE:
  NAME/KEY: CDS
  LOCATION: join(713..721, 981..1250)
  US-08-445-265A-1

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250      Gaps: 0
  Percent Similarity: 88.889      Percent Identity: 66.667

alignment_block:
  US-09-471-276-831_COPY_1_16 x US-08-445-265A-1  ..
Align seg 1/1  to: US-08-445-265A-1  from: 1  to: 4283
      8 leu1e0u1e0t7p1c3yvaltrh1t7p1c3y 16

```

```

      ::::::::::::::::::::::::::::
      1480 ATTCTATTCTGCGGGGTGCGGTGCGGCGC 1506

seq_name: /cgn2_6/p1c0data/2/1na/6A.COMB.seq:US-08-990-442-1
seq_documentation_block:
  Sequence 1, Application US/08990442
  Patient No. 6090790
  GENERAL INFORMATION:
  APPLICANT: Eriksson, Elof
  TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
  NUMBER OF SEQUENCES: 4
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Charles A Brady
  STREET: 1 South Pritchney Street
  CITY: Madison
  STATE: WI
  COUNTRY: US
  ZIP: 53703
  COMPUTER READABLE FORM:
  MEDIUM TYPE: floppy disk
  OPERATING SYSTEM: IBM PC COMPATIBLE
  SOFTWARE: PC-DOS/MS-DOS
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/990.442
  FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
  NAME: Berison, Bennett J
  REGISTRATION NUMBER: 37094
  REFERENCE/DOCKET NUMBER: 310558.90028
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 608-251-5000
  TELEFAX: 608-251-9166
  INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
  TYPE: 4283 base pairs
  STRANDEDNESS: double
  TOPOLOGY: circular
  MOLECULE TYPE: other nucleic acid
  FEATURE:
  NAME/KEY: CDS
  LOCATION: join(713..721, 981..1250)
  US-08-990-442-1

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250      Gaps: 0
  Percent Similarity: 88.889      Percent Identity: 66.667

alignment_block:
  US-09-471-276-831_COPY_1_16 x US-08-990-442-1  ..
Align seg 1/1  to: US-08-990-442-1  from: 1  to: 4283
      8 leu1e0u1e0t7p1c3yvaltrh1t7p1c3y 16
      ::::::::::::::::::::::::::::
      1480 ATTCTATTCTGCGGGGTGCGGTGCGGCGC 1506

seq_documentation_block:
  Sequence 1, Application US/08760615
  Patient No. 6090790
  GENERAL INFORMATION:
  APPLICANT: Haynes, Joel R
  APPLICANT: Schmaljohn, Connie S
  APPLICANT: Fuller, Deborah L
  APPLICANT: Schmaljohn, Alan

```

```

Patent No.6197332
GENERAL INFORMATION:
APPLICANT: Ronald Zuckermann et al.
TITLE OF INVENTION: Lipid-conjugated polyanide Compounds and Related
ATTC NO.: 1711
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Chilton Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
SEQUENCE NO.: 1
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcutiln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132-808
PRIORITY:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 187,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 523-5476
FAX: (510) 523-5476
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4328 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-132-808-1

alignment_scores:
Quality: 42.00 Length: 9
Ratio: 5.250 Gaps: 0
Percent Similarity: 88.889 Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-09-132-808-1 ..

Align seq 1/1 to: US-09-132-808-1 from: 1 to: 4328

      8 LeuLeuLeuTrpGlyAlaThrTrpGly 16
      :::::::::::
1862 ATCTTATCTTGGGAGGAGGAGGAGGAGGAGG 1888

seq_name : cpg2_6/cfcd42/1na/5B_Cbma.seq:US-08-910-647-2

seq_documentation_block:
Sequence 2, Application US/08910647
Patent No. 6251433
GENERAL INFORMATION:
APPLICANT: Ronald Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
TITLE OR INVENTION: Polynucleotide Delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Chilton Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
SEQUENCE NO.: 1
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcutiln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132-808
PRIORITY:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 187,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 523-5476
FAX: (510) 523-5476
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4328 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-132-808-1

```

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/910,647
? PUBLICATION NUMBER:
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Fujita, Sharon M.
? REGISTRATION NUMBER: 38,459
? TELECOMMUNICATION INFORMATION:
? REFERENCE/DOCKET NUMBER: 1218.002
? TELEPHONE: (510) 655-3542
? TELEFAX: (510) 655-3542
? INFORMATION FOR SEQ ID NO.: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4328 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-910-647-2

alignment_scores:
? Quality: 42.00 Length: 9
? Ratio: 5.250 Gaps: 0
? Percent Similarity: 88.889 Percent Identity: 66.667
? alignment_block:
? US-09-471-276-831_COPY_1_16 x US-08-910-647-2 ..
?
? Align seq 1/1 to: US-08-910-647-2 from: 1 to: 4328
?
? 8 LeuLeuLeuTTCGlyValThrTrpGly 16
? ..:::|||||:::|||||:::|||||
? 1862 ATTCATTCTCGGGGGGTGGGGGGGGC 1888
?
? seq_name: /cgn2_5/pdata2/1/mn/5A.COMB.seq:US-08-910-647-4
?
? seq_documentation_block:
? Sequence 4, Application US/08910647
? Patent No. 6251433
? GENERAL INFORMATION:
? INVENTOR: Peter H. Rabinowitz et al
? INVENTOR ADDRESS: 10000 Wilshire Blvd, Suite 200
? TITLE OF INVENTION: Compositions and Methods for
? Polynucleotide Delivery
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Chiron Corporation
? STREET: 4560 Horton Street
? CITY: Emeryville
? STATE: CA 94608
? COUNTRY: CA 94608
? ZIP: 94608-2916
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/910,647
? FILING DATE:
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Fujita, Sharon M.
? REGISTRATION NUMBER: 38,459
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 923-2705
? TELEFAX: (510) 655-3542
? INFORMATION FOR SEQ ID NO.: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4818 base pairs

```

```

? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-910-647-4

alignment_scores:
? Quality: 42.00 Length: 9
? Ratio: 5.250 Gaps: 0
? Percent Similarity: 88.889 Percent Identity: 66.667
? alignment_block:
? US-09-471-276-831_COPY_1_16 x US-08-910-647-4 ..
?
? Align seq 1/1 to: US-08-910-647-4 from: 1 to: 4818
?
? 8 LeuLeuLeuTTCGlyValThrTrpGly 16
? ..:|||||:::|||||:::|||||
? 2354 ATTCATTCTCGGGGGGTGGGGGGGGC 2380
?
? seq_name: /cgn2_5/pdata2/1/mn/5A.COMB.seq:US-08-345-913-1
?
? seq_documentation_block:
? Sequence 1, Application US/08345913
? Patent No. 5641665
? GENERAL INFORMATION:
? APPLICANT: Hoechst, Peter
? INVENTOR: Peter H. Rabinowitz et al
? APPLICANT ADDRESS: Merck & Co., Inc.
? APPLICANT: Rahldbl, Shilin
? TITLE OF INVENTION: PLASMIDS SUITABLE FOR IL-2 EXPRESSION
? NUMBER OF SEQUENCES: 1
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knudsen, Martens, Olson and Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: CA
? COUNTRY: USA
? ZIP: 92660
? COMPUTER READABLE FORM:
? MEDIUM TYPE: diskette
? OPERATING SYSTEM: IBM compatible
? SOFTWARE: PatentIn Release 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/345,913
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? ATTORNEY/AGENT INFORMATION:
? NAME: Mays, Vensko, Nancy
? REGISTRATION NUMBER: 36,298
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-235-8550
? TELEFAX: 619-235-0176
? INFORMATION FOR SEQ ID NO.: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4928 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOCHROMIC: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE:
? ORIGINAL SOURCE:
? FEATURE:
? NAME/KEY: Coding Sequence

```

```

; LOCATION: 1689...2159
; OTHER INFORMATION:
US-08-345-913-1

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250        Gaps: 0
  Percent Similarity: 88.889   Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-345-913-1 ..
Align seg 1/1 to: US-08-345-913-1 from: 1 to: 4928

8 LeuLeuLeuTPFGlyValThrTPFGly 16
:::|||||:|||||:|||||:
2388 ATTCAATTCCTGGGGGTGGGGTGGCGGC 2414

seq_name: /cgn2_6/p/cdata/2/lna/58.COMB.seq:US-08-818-562-1
seq_documentation_block:
; Sequence 1, Application US/08818562
; Patent No. 6147055
; GENERAL INFORMATION:
; APPLICANT: Robert, Peter W.
; APPLICANT: Parke-Davis, Inc.
; APPLICANT: Parke-Davis, Inc. E.
; TITLE OF INVENTION: Plasmids suitable for IL-2 Expression
; FILE REFERENCE: 1530, 0080001
; CURRENT APPLICATION NUMBER: US/08/818,562
; CURRENT FILING DATE: 1997-03-14
; PUBLICATION NUMBER: US 98/345,913
; EARLIER APPLICATION NUMBER: US 94-411,28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentia Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4928
; TYPE: DNA
; ORGANISM: Homo sapiens
; GENE: IL2
; NAME/KEY: CDS
; LOCATION: (1689)..(2159)
US-08-818-562-1

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250        Gaps: 0
  Percent Similarity: 88.889   Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-818-562-1 ..
Align seg 1/1 to: US-08-818-562-1 from: 1 to: 4928

8 LeuLeuLeuTPFGlyValThrTPFGly 16
:::|||||:|||||:|||||:
2388 ATTCAATTCCTGGGGGTGGGGTGGCGGC 2414

seq_name: /cgn2_6/p/cdata/2/lna/58.COMB.seq:US-08-564-313-1
seq_documentation_block:
; Sequence 1, Application US/08564313
; Patent No. 6147055
; GENERAL INFORMATION:
; APPLICANT: Nabel, Elizabeth
; APPLICANT: Nabel, Gary
; APPLICANT: Lew, Denise
; APPLICANT: Marquez, Magda
; TITLE OF INVENTION: PLASMIDS SUITABLE FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knoble, Martens, Olson and Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,313
; FILING DATE: 01-DEC-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/074,344
; FILING DATE: 07-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: USM150000
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: VICAL 033CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4965 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA
; ANTI-SENSE: NO
; FRAGMENT TYPE: NO
; IMMEDIATE SOURCE:
; CLONE: HLA-B7 and Beta-2
US-08-564-313-1

alignment_scores:
  Quality: 42.00      Length: 9
  Ratio: 5.250        Gaps: 0
  Percent Similarity: 88.889   Percent Identity: 66.667

alignment_block:
US-09-471-276-831_COPY_1_16 x US-08-564-313-1 ..
Align seg 1/1 to: US-08-564-313-1 from: 1 to: 4965

8 LeuLeuLeuTPFGlyValThrTPFGly 16
:::|||||:|||||:|||||:
3021 ATTCAATTCCTGGGGGTGGGGTGGCGGC 3047

seq_name: /cgn2_6/p/cdata/2/lna/PCTUS.COMB.seq:PCT-US94-06065-1
seq_documentation_block:
; Sequence 1, Application PC/US9406069
; GENERAL INFORMATION:
; APPLICANT: Vical Incorporated
; APPLICANT: Regents of the University of Michigan
; APPLICANT: Nabel, Elizabeth
; APPLICANT: Nabel, Gary
; APPLICANT: Lew, Denise
; APPLICANT: Marquez, Magda
; TITLE OF INVENTION: PLASMIDS SUITABLE FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knoble, Martens, Olson and Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach

```



```

STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: FASTSD Version 1.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06069
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: 07/JUN/1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET INFORMATION: VICAL, 033VPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4965 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
STRAIN: HLA-B7 and Beta-2
PCT-US94-06069-1

```

Alignment scores:

Quality:	42.00	Length:	9
Ratio:	5.250	Gaps:	0
Percent Similarity:	88.889	Percent Identity:	66.667

Alignment block:

```

US-09-471-276-831_COPY_1_16 x PCT-US94-06069-1
Align seq 1/1 to: PCT-US94-06069-1 from: 1 to: 4965

```

8 LeuLeuLeuTrpGlyValThrTrpIly 16
 ::::::::::::::::::::::::::::

seq.name: /cgn2_6/pcldata/2/lna/68_CONB.seq:US-08-910-647-3

seq.document: block:

```

Sequence 3, Application US/08910647
Patent No. 6251433
GENERAL INFORMATION:
APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and methods for
TITLE OF INVENTION: Polynucleotide delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: United States of America
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,647
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET INFORMATION: 1218,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5107 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-910-647-3

```

Alignment scores:

Quality:	42.00	Length:	9
Ratio:	5.250	Gaps:	0
Percent Similarity:	88.889	Percent Identity:	66.667

Alignment block:

```

US-09-471-276-831_COPY_1_16 x US-08-910-647-3
Align seq 1/1 to: US-08-910-647-3 from: 1 to: 5107

```

8 LeuLeuLeuTrpGlyValThrTrpIly 16
 ::::::::::::::::::::::::::::

2643 ATCTATTCTCGGGCGCTGGCGGC 2659

seq.name: /cgn2_6/pcldata/2/lna/58_CONB.seq:US-08-073-836-3

seq.document: block:

```

Sequence 3, Application US/08073836
Patent No. 5650306
GENERAL INFORMATION:
APPLICANT: Nabel, Gary J.
APPLICANT: Yang, Zhi yong
TITLE OF INVENTION: RECOMBINANT NUCLEIC ACIDS FOR INHIBITING
TITLE OF INVENTION: HIV GENE EXPRESSION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,836
FILING DATE: 07-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET INFORMATION: 1218,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

```

```

?      LENGTH: 5653 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: double
?      ORIENTATION: linear
US-08-073-836-3

```

```

alignment_scores:
    Quality: 42.00      Length: 9
    Ratio: 5.250      Gaps: 0
    Percent Similarity: 88.869      Percent Identity: 66.667

```

alignment_block:

US-09-471-276-831_copy_1_16 x US-08-073-836-3 ...

Align seg 1/1 to: US-08-073-836-3 from: 1 to: 5653

8 leuLeuLeuTrpGlyValTrpTrpCly 16

1286 ATCTATTCTCGGGGGGCTCGCTCGCGC 1312

seq_name: /cgn2_6/prodata/2/lna/5A_Comb.seq:US-08-235-277-1

seq_documentation_block:

Sequence 1, Application US/08235277

Patent No. 5733543

INVENTOR: GARY J

APPLICANT: WOFFENDIN, CLIVE

APPLICANT: YANQ, NIN-SUN

APPLICANT: SHEEHY, MICHAEL J

TITLE OF INVENTION: INTRODUCTION OF HIV-PROTECTIVE GENES

TITLE OF INVENTION: INFO CELLS BY PARTICLE-MEDIATED GENE TRANSFER

NUMBER OF SEQUENCES: 1

ADDRESSER: P.C. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

COMPUTER: IBM PC COMPATIBLE

MEDIUM TYPE: FLOPPY DISK

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/235,277

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5733543man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 6042-008-68

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-3200

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5653 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

ORIENTATION: unknown

MOLWT: 1286

OTHER nucleic acid

US-08-235-277-1

alignment_block:

US-09-471-276-831_copy_1_16 x US-08-235-277-1 ...

Align seg 1/1 to: US-08-235-277-1 from: 1 to: 5653

8 leuLeuLeuTrpGlyValTrpTrpCly 16

1286 ATCTATTCTCGGGGGGCTCGCTCGCGC 1312

```

alignment_scores:
    Quality: 42.00      Length: 9
    Ratio: 5.250      Gaps: 0
    Percent Similarity: 88.869      Percent Identity: 66.667

```



```

? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
? COMMENT: FOR READABLE FORM:
? MODIFICATION: 1997-03-01
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentia Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/985,950
? FILING DATE: 05-DEC-1997
? PRIORITY: 05
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/041,279
? FILING DATE: 21-MARCH-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/033,181
? FILING DATE: 16-DEC-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/032,252
? FILING DATE: 05-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Ching, Edwin P.
? REGISTRATION NUMBER: 34,090
? REFERENCE/DOCKET NUMBER: DX0670K
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 953,1296
? TELEFAX: (650) 953,1240
? INFORMATION FOR SEQ ID NO: 21:
? SEQUENCE CHARACTERISTICS: 21:
? LENGTH: 2790 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 177..2132
? NAME/KEY: misc_feature
? LOCATION: 1772
? OTHER INFORMATION: /note= "nucleotide 1722 designated
US-08-985-950-21"

Alignment_scores:
Quality: 98.50 Length: 121
Ratio: 1.449 Gaps: 7
Percent Similarity: 56.198 Percent Identity: 31.864

Alignment_block:
us-09-471-276-831 x US-08-985-950-21 ..
Align seq 1/1 to: US-08-985-950-21 from: 1 to: 2790

1 McSetMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGlyPrt 17
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
177 ATGACCCCGCTCTCTACAGGCTGATGCTGTCTGGAGCTGAGCTGAGCC 226
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
17 oValThrGlnAlaIlePheTrpGlnThrGln***SerLeuTrpAlaG 34
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
227 CCGAGCCCGCTGACGAGCGACGCTCCCAAGCCACCCCTCTGCGCGTG 376
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
34 IuSerGluHis**LeuYsThrLeuGly.GInGlyAspAlaAspValP 50
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
277 AACCAAGCTCTGTGATC...ACCAGGAGAGCTCTGTGACCTGAGCTG 323
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
50 rodGlyProProGlyAspSerArgLeuProAlaVal.GInGlu..... 63
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
324 CAGGCGGCGCAGAGACCGAGGAGTACCTGTATATAGAGAAAGAAAGAC 373
64 .....TTPGlyAlaGlnGluProValHisLeuAspSerProAlaIleL 78

```

```

374 AGCACCTGATATACAGATCCACAGACACTT.....GTGAAGA 414
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
78 YshISgInPhe.....LeuLeuThrGlyAspThrGlnGlyArgTr 91
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
415 AGGCGCATTTCCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 464
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
92 ArgGysArgSerGlyLeuSerTrp...GlyTrp**GlnIuSerYsLe 107
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
465 CCTCTTACTATGCTATAGCAGACACTGAGCGCGCTCATAGAGACAGTACC 514
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
107 ULaGlnGluLeu 110
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
515 CCTGCAAGCTG 524
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

seq_name: /cgn2_6/prodata/2/lna/68_Comb.seq:us-09-345-468-24

seq_documentation_block:
: Sequence 24, Application US/09345468
: Patent No. 6243527
: ORIGINATOR: UNKNOWN
: APPLICANT: Villaveal, S.
: APPLICANT: Villaveal, J.
: APPLICANT: Jandrot-Perrus, M.
: TITLE OF INVENTION: GLYCOPROTEIN VI AND USBS THEREOF
: FILE REFERENCE: 7853-147
: CURRENT APPLICATION NUMBER: US/09/345,468
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 24
: LENGTH: 1896
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-345-468-24

Alignment_scores:
Quality: 87.50 Length: 138
Ratio: 1.509 Gaps: 6
Percent Similarity: 42.029 Percent Identity: 28.261

Alignment_block:
us-09-471-276-831 x US-09-345-468-24 ..
Align seq 1/1 to: US-09-345-468-24 from: 1 to: 1896

1 McSetMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGlyPrt 17
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
1 ATGACCGCGCGCTCTCTACAGCCCTGCTCTGCTCTGCTGCTGAGCTGAGCC 50
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
17 oValThrGlnAlaIlePheTrpGlnThrGln***SerLeuTrpAlaG 34
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
51 CAGCAGCCCGCTGACGAGCGCTCTCCCAAGCCACCCCTCTGCGCGTG 100
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
34 IuSerGluHis**LeuYsThrLeuGlyGInGlyAspAlaAspValPro 50
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
101 AG.....AlaGlnGlu..... 68
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
51 GlyProProGlyAspSerArgLeuProAlaVal.GInGluTrpGly..... 65
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
103 .....CCAGCC.....TCTGTATACAGCTGGGGGAGCC 131
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
66 .....AlaGlnGlu..... 68
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
132 CGTGACCATCTGTGTCAGAGGAGCGCTGAGAGCCAGAGAGTACCACTGC 181
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
69 .....ProValHisLeuAsp.....SerPro 75
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
182 ATAAAGAGGAGAAAGCCAGACCCCTTGGACAGAAATTAACCCCACTGAGACC 231
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
76 AlaIleYshISgInPheLeuLeuThrGlyAspThrGln.....G1 89

```



```

81 ATGACCCCACTGCTACGGTCTGATCTGTCTGGGCTGAGCC 130
17 CvalThrcGluAlaIlePheTyrgLutHrgIn***SeIeUTrPAlaG 34
131 CAGGACACAGCGACAGGAGCCCTCCGACAGCCAGCCCTCGAGCTG 180
34 IuSerGluHls***LeuIyThrlLeuGly..GInCyAspAlaIaPAlP 50
181 AGCCAGGCTGTGTATC...ACCCAGGAGAGTCTGTGGACCTCGACGCT 227
50 TcdIyPrProGlyAspSerArgLeuPProAlaIval..GInGlu..... 63
228 CAGGAGACCTCGAGACAGACAGCTGCTGATCTGTATGAGAAAGAAAC 277
64 .....TrGcIyAlaGInGluPProValHlsLeuAspSerProAlaIleL 78
278 AGCACTCGATTAACAGGATCCGACAGGACTT.....GTGAGAG 318
78 yHslGInPhe.....LeuLeuThrcGlyAspThrcGInGlyTyr 91
319 AGGCGACATTCGACATCTGATCACTACGCTGGAGACAGCGAGGAG 368
92 ArgCyAsrSerGlyLeuSerThr...GlyTyr**+InLeuSerIyLe 107
369 TGCTGTATCTATGCGACGACACTGCGAGGCTCTCGAGAGAGTCCAC 418
107 UdaGcIuLeu 110
419 CTGTGAGCTG 428

```

seq_name: /cgn2_6/prodata/2/lna/fa_COMB.seq:US-09-135-782-3

seq_documentation_block:

```

: Sequence 3, Application US/09135782
: GENERAL INFORMATION:
: APPLICANT: Xu, Shuang-yong
: TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction
: TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The
: FILE REFERENCE: M88-143
: CURRENT APPLICATION NUMBER: US/09/135,782
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 3
: LENGTH: 735
: TYPE: DNA
: ORGANISM: No. 6027929loc sp.
: NAME/REV: CDS
: LOCATION: (1)..(732)
US-09-135-782-3

```

allignment_scores:

```

Quality: 78.50 Length: 56
Size: 2092
Percent Similarity: 64.286 Percent Identity: 59.286

```

allignment_block:

US-09-471-276-831 x US-09-135-782-3 ..

Align seg 1/1 to: US-09-135-782-3 from: 1 to: 735

```

55 AspSerArgLeuPProAlaIvalGInGluPProValHlsLeuAspSerProAlaIleL 71
|||||.....|
394 CATCATGATTTCG...ATCGAAGATGGGAGCTGAGAGGCTTCACAA 440
71 sLeuAspSer.....ProAlaIleIyHslGInPhe...LeuLeuT 84
|||||.....|
441 CTTAACATCTATGATATGATGAAATACAGCAAGATTCCTCAATTGTGA 490

```

```

84 hrcIyAspThrcGInGlyArgTyr.....CysArg 94
|||||.....|
491 AGAAGATACCTCGGACATCAATCTTTAAATCAGAGAGCTGTAC 540
95 SerGlyLeuSerThrcIy 100
541 CGATGTATCAAACTGCT 558

```

seq_name: /cgn2_6/prodata/2/lna/fa_COMB.seq:US-09-172-108-48

seq_documentation_block:

```

: Sequence 48, Application US/09172108
: GENERAL INFORMATION:
: APPLICANT: Cunningham, Mary Jane
: APPLICANT: Zwiesler, Gary B.
: APPLICANT: Panzer, Scott R.
: APPLICANT: Sellhammer, Jeffrey J.
: TITLE OF INVENTION: MARKERS FOR PEROXISOMAL PROLIFERATORS
: FILE REFERENCE: PA-0012 US
: CURRENT FILING DATE: 1998-10-13
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PERL Program
: SEQ ID NO: 48
: LENGTH: 285
: TYPE: DNA
: ORGANISM: homo sapiens
: NAME/REV:
: OTHER INFORMATION: 700607235H1
US-09-172-108-48

```

allignment_scores:

```

Quality: 74.50 Length: 65
Size: 1833
Percent Similarity: 61.538 Percent Identity: 56.923

```

allignment_block:

US-09-471-276-831 x US-09-172-108-48 ..

Align seg 1/1 to: US-09-172-108-48 from: 1 to: 285

```

1 MetSerLeuLeuValPheLeuLeuThrcGlyValThrTrGcIyPr 17
|||||.....|
14 ATCTCTCTGCTGACTACGTACTTCTCTGGGCTTATTATTGGGCCC 63
17 CvalThrcGluAlaIlePheTyrgLutHrgIn***SeIeUTrPAlaG 34
|||||.....|
64 AGCACTACGACAGCGCTGATATTT.....CAAGAGAGCTGGAGAAAC 106
34 IuSerGluHls***LeuIyThrlLeuGlyGInCyAspAla..... 47
|||||.....|
107 AGTCC.....CTTGCCAGGCGCTGGCTTTCGCAATCCATCGAGCCCTG 150
48 .....AspValProGlyPrProGlyAspSer 56
131 CATCACTGCTGCTGACGACATCTGCTGCTCCAGCGAGAGCT 195

```

seq_name: /cgn2_6/prodata/2/lna/fa_COMB.seq:US-09-172-711-45

seq_documentation_block:

```

: Sequence 45, Application US/09172711
: Patent No. 6160105
: GENERAL INFORMATION:
: APPLICANT: Cunningham, Mary Jane
: APPLICANT: Zwiesler, Gary B.
: APPLICANT: Panzer, Scott R.
: APPLICANT: Sellhammer, Jeffrey J.
: TITLE OF INVENTION: MONITORING TOXICOLOGICAL RESPONSES
: FILE REFERENCE: PA-0011 US
: CURRENT FILING DATE: 1998-10-13

```

[illegible]

```
seq_documentation_block:
; Sequence 1, Application US/08642541
; Patent No. 5891704
```

```

2  SerLeuLeuVal1AlaPheLeuLeuIleProValIah.....TTP.G.16
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
691 GCATGGCAACATCACTGCTGGCATTCGTGCTGGCACCAACAGCATGGT 740
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
16  IYProValIhGcIuaIaIaIleIheThyTgIuThGIn**SerLeuTTP.32
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
741 GAGCATATACCTCTGACGAGGCTCTAGGACATCAAGCGTGGTGGTCTTC 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33  AlAluSerCuiIuIis**IeulYsThrIeulCylIcYcAspAlaIayA.4
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
791 GCGCCATGACGCCGCACTTGTATATGCGCCATCAACAGCTCACTCTG.840
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49  IProGlyIProPro.....GlyAspSerArgIeuProAlay.61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
841 GCATGAGCGCCCACTGCCCAACGCTCAAGCTCTGGCGAATTCCTCGAC.890
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61  aIeInGluTrfGlyIaIeInGluPro.....
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
891 CGGAGGCGCAGGGGTGAGCTATATCTTACCGCGAGCTGGCGAGCTTCC.940

```



```

70. ValHis.....LeuAspSerProValAlaIleHisGln 81
      |||||.....|
941 GAGTACACCCCTGGCCCGCCAGCAGACATGAGCCGCGGCGATGATCG 960
      |||.....|
81 GAGTAACTTCTATGATGTTTCTGATCTATGATGATGATGATGATGATG 97
      |||.....|
991 CCTT.....CGAATCTCAAGGCGCGCATCTGTCGCGGCGGGT 1030

seq_name: /cogn2_6/prodbeta2/lna/75L.COMB.seq:us-08-642-541-3

seq_documentation_block:
1 Sequence 3 application US/08642541
2 Accession 589730
3 GENERAL INFORMATION:
4 APPLICANT: TUYING, TAN
5 TITLE OF INVENTION: HIGH EXPRESSION MODULES ENCODING
6 TITLE OF INVENTION: METHIONINASE, PURIFIED RECOMBINANT METHI
7 TITLE OF INVENTION: THEREFROM AND USES THEREOF
8 NUMBER OF SEQUENCES: 8
9 CORRESPONDENCE ADDRESS:
10 ATTORNEY/AGENT INFORMATION:
11 STREET: 2000 Peninsula Avenue, NW
12 CITY: Washington
13 COUNTRY: USA
14 ZIP: 20006-1888
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: floppy disk
17 SOFTWARE: PC-DOS/MS-DOS
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.30
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/642,541
22 FILING DATE: 03-MAY-1996
23 CLASSIFICATION: 424
24 ATTORNEY/AGENT INFORMATION:
25 STREET:
26 CITY:
27 COUNTRY:
28 REGISTRATION NUMBER: 36,217
29 REFERENCE/DOCKET NUMBER: 3176-20002, 21
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (202) 887-1500
32 TELEFAX: (202) 822-0168
33 TELEX: 90-4030
34 INFORMATION FOR SEQ ID NO: 3:
35 SOURCE: CHINA PATENT
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 1369 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: double
40 TOPOLOGY: linear
41 MOLECULE TYPE: DNA (genomic)
42 ANTI-SENSE: YES
43 US-08-642-541-3

alignment_scores:
      Quality: 74.50      Length: 117
      Ratio: 1.242      Gaps: 6
Percent Identity: 51.262      Percent Identity: 29.060

alignmen_block:
US-09-411-216-631 x US-08-642-541-3/rev ..

Align seg 1/1 to reverse of: US-08-642-541-3 from: 1 to: 1369

      2 SerMetLeuValValaPheIleuLeuLeuTPrGValaTTr.....TPr G 16
      |||:|||||.....|
679 GCGATCGCGACCTGATCTGCTGCTGCTGCTGCGGCGCGATGCTG 630
      |||:|||||.....|
16 IPrGValaTTrGValaAlaAlaIlePheTTrGValaTTr.....SerLeuTTr 32
      |||:|||||.....|
629 GACCTTGTATCTGTGACGAGCGCTCAAGACATGACGCGGTGCTGTC 580
      |||:|||||.....|
33 AlaGluSerGluHis***LeuValaTTrLeuAlaTTrGValaTTrGVala 49

```

```

579 GCGCCATGACCGCCAGCTTGATGATCGCGCATACACCTTCACCTC 530
      |||
49 IPDgIyProHo.....GIyAspSerArgIleProAlaIy 61
      |||
539 GAGTGGACCGCCATCGCCCGACACCTGACAGTGGCGCAAGTCTCGCC 480
      |||
61 aIeInguIyArgIAlaInguIyPro..... 69
      |||
479 CGCAGCGCCCGACGAGTGGAGTATCATTCATTCACCGCGCTTCGCCACCTTCC 430
      |||
70 ..yAlHis.....LeuAspSerProAlaIleIyHisIleInp 81
      |||
429 GAGATACACCTTGGCCCGCCGACACGATAGCAGCGCGCGGATGATGTC 380
      |||
81 eIeuleIeHfIyAspHfInGIyArgIyArgGyArgSerG 370
      |||
379 CCTT.....CGACATGAGGGGGGCGATCGTGGCGGCGCGCGT 340
      |||

seq_documentation_block:
? Sequence 1, Application US/09260889
? Patent No. 6231854
? GENERAL INFORMATION:
? ALTERNATIVE: NM, IyInp
? TITLE OF INVENTION: NM EXPRESSION MODULES ENCODING
? NUMBER OF SEQUENCES: 8
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MORRISON & FOENSTER
? STREET: 2000 Pennsylvania Avenue, NW, suite 5500
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? PUBLICATION NO.: 20060889
? COMPUTER READER: COM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: Windows
? SOFTWARE: FASTSD for Windows Version 2.0b
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/760,889
? FILING DATE:
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? PRIOR APPLICATION NO.: 08/914,377
? FILING DATE: 19-AUG-1997
? APPLICATION NUMBER: 08/662,541
? FILING DATE: 03-MAY-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Kurashige, Kate H
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET INFORMATION: 31276-20002.11
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-8601300
? TELEFAX: 202-822-0189
? TELETYPE:
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1369 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: Linear
? MAPPING TYPE: Genomic DNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 48...1241
? OTHER INFORMATION:
? US-09-260-889-1

```

Alignment scores:
 Quality: 74.50 Length: 117
 Ratio: 1.242 Gaps: 6
 Percent Similarity: 51.282 Percent Identity: 29.060

Alignment block:
 US-09-471-276-831 x US-09-260-889-1 ..

Align seg 1/1 to: US-09-260-889-1 from: 1 to: 1369

```

2 SerMetLeuValValPheLeuLeuLeuTrpLysAlaThr.....TTP G 16
591 GCGATCGGCGACATGCTGCTGCATGCTGCTGCGGACGACGACGCTGCTG 740
16 1yProValThrGluAlaIlePheTyrGluThrGln***SerLeuTrp 32
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
741 GACCGTATGCTGCTGAGGCGCTGACAGGACATGACGCGTGGCGTCTG 790
33 AluGluSerGluHis***LeuLysThrLeuLysGlnGlyIncysAspAlaAspVa 49
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
791 GCGCCATGACGCGCGACGCTTGTATGCGCGGATGACACCCGTCACACGTCG 840
49 1ProGlyProPro.....GlyAspSerArgLeuProAlaVal 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
841 GCATGACGCGCGACGCTGCTGACGCTGACGCTGCTGCTGCTGCTGCTG 890
61 AluGluTrpTrpLysAlaGlnGluPro.....SerLeuTrp 69
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
891 GCGGACGCGCGACGCTGACGCTATTCATCCGCGCGCTGCGACGCTGCGC 940
70 ...ValHis.....LeuAspSerProAlaIleLysHisGlnPhe 81
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
941 GCGATGACGCTGCGCGCGCGACGACGATGACGACGCGCGCGCTATGTC 990
81 eLeuLeuThrGlyAspThrGlnGlyArgTyrArgGlyAspGlyLeu 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 CCTT.....CGACGTCACAGCGCGCGCATCGTGGCGCGCGCGCTT 1030

```

seq_name: /cgn2.6/prodata/2/lna/5a.COWB.seq US-09-260-889-3

seq_documentation_block:

Sequence 3 Application US/09260889

Patent No. 6231854

GENERAL INFORMATION:

APPLICANT: TAN, Yuying

TITLE OF INVENTION: HIGH EXPRESSION MODULES ENCODING

NUMBER OF SEQUENCES: 8

ADDRESS: MORRISON & FORESTER

STREET: 2000 Pennsylvania Avenue, NW, suite 5500

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

EDITED BY: J. H. H. H.

COMPUTER: IBM Compatible

SOFTWARE: PASTED FOR Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/260.889

FILING DATE:

CLASSIFICATION: 424

PROVISIONAL PATENT:

APPLICATION NUMBER: 08/514.377

FILING DATE: 19-AUG-1997

APPLICATION NUMBER: 08/642.541

FILING DATE: 03-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mursallig, Kate H

REGISTRATION NUMBER: 29, 959

REFERENCE/DOCKET NUMBER: 312/6-20002.11

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-887-1500

TELEFAX: 202-822-0168

INSTRUMENT FOR SEQ. ID. NO. 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1369 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: Linear

MOLECULE TYPE: Genomic DNA

REFERENCE: YES

US-09-260-889-3

Alignment scores:
 Quality: 74.50 Length: 117
 Ratio: 1.242 Gaps: 6
 Percent Similarity: 51.282 Percent Identity: 29.060

Alignment block:
 US-09-471-276-831 x US-09-260-889-3/rev ..

Align seg 1/1 to reverse of: US-09-260-889-3 from: 1 to: 1369

```

2 SerMetLeuValValPheLeuLeuLeuTrpLysAlaThr.....TTP G 16
679 GCGATGCGCGACATGCTGCTGCATGCTGCTGCGGACGACGACGCTGCTG 630
16 1yProValThrGluAlaIlePheTyrGluThrGln***SerLeuTrp 32
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
629 GACCGTATGCTGCTGAGGCGCTGACAGGACATGACGCGTGGCGTCTG 580
33 AluGluSerGluHis***LeuLysThrLeuLysGlnGlyIncysAspAlaAspVa 49
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
579 GCGCCATGACGCGCGACGCTTGTATGCGCGGATGACACCCGTCACACGTCG 530
49 1ProGlyProPro.....GlyAspSerArgLeuProAlaVal 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
529 GCATGACGCGCGACGCTGCTGACGCTGACGCTGCTGCTGCTGCTGCTG 480
61 AluGluTrpTrpLysAlaGlnGluPro.....SerLeuTrp 69
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
479 GCGGACGCGCGACGCTGACGCTATTCATCCGCGCGCTGCGACGCTGCGC 430
70 ...ValHis.....LeuAspSerProAlaIleLysHisGlnPhe 81
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 GCGATGACGCTGCGCGCGCGACGACGATGACGACGCGCGCGCTATGTC 380
81 eLeuLeuThrGlyAspThrGlnGlyArgTyrArgGlyAspGlyLeu 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 CCTT.....CGACGTCACAGCGCGCATCGTGGCGCGCGCGCTT 340

```

seq_name: /cgn2.6/prodata/2/lna/5a.COWB.seq US-08-384-106A-8

seq_documentation_block:

Sequence 8 Application US/09384106A

Patent No. 6231854

GENERAL INFORMATION:

APPLICANT: Sheff Ph.D., Charles J.

APPLICANT: Downing M.D., James J.

APPLICANT: HIRAI Ph.D., Hiroshi

APPLICANT: Okada, Tsukasa

TITLE OF INVENTION: In4c-p18 and In4d-p19, Inhibitors of

NUMBER OF SEQUENCES: 25

ADDRESS: STEPHEN KESSLER, GOLDSTEIN & FOX

STREET: 1100 New York Ave., N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

```

COMPUTER READABLE FORM:
MEDICIN TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,106A
FILING DATE: 06-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0556 0500000
TELEPHONE: 202-371-2500
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
TRANSDUCED: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..501
US-08-384-106A-8

alignment_scores:
Quality: 74.00 Length: 99
Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303

alignment_block:
US-09-471-276-831 x US-08-384-106A-8 ..
Align seq 1/1 to: US-08-384-106A-8 from: 1 to: 501

17 ProvalThrluaAlaIaIle.....PhetygluThrcIn**Serle 31
||||| :||||| ||| :|||||
229 CCAGTCATCAGCAGCCGCCGCTGCTGCGACAC..... 287
31 utPalaIuSerGIuHs***LeuYshrlaenGIyInCysaSPAlaA 48
spvalProlGIyPrProGIyAspSerArGIaPro.....AlaVal 61
268 .....CTGAAGTCCTAGCGACACGGGCGTC 295
48 spvalProlGIyPrProGIyAspSerArGIaPro.....AlaVal 61
||||| ||| ||| :|||||
296 ATGCACAGTGCCTGATGGACACGGGCGCTTCATCTCGACGCTT 345
62 GInGlu.....TrpGlyAlaGInGluProva 70
||||| :|||||
346 CAGAGGCGTCACTGCTGTGTCAGCTTTCGACGCTGATATCATCT 395
70 lHsLeuaAspSerProAlaIleYshrlaenGInPheLeuThrGIyasp 86
||||| :||||| ||| :|||||
396 CCATCGACGGAGACCGCGGCGTCTACACCCCTTG..... 430
87 thGInGInGIyArGIYArGIYAspSerGIyLeuSerThrGIyTPR 101
||||| ||| ||| :|||||
431 .....ACGTGGACGTCGAGACGCGCTCAGACCTCGTGG 466

seq_name: /cgn2.6/plodata/2/1na/PCTUS.COMB.seq:PCT-US96-01643-8
seq_documentation_block:
SEQUENCE INFORMATION PC/70US9601643
GENERAL INFORMATION
APPLICANT: St. Jude Children's Research Hospital
TITLE OF INVENTION: Inhibitors of
TITLE OF INVENTION: Inhibitors of
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
COMPUTER READABLE FORM:
MEDICIN TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01643
FILING DATE: 06-FEB-1996
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/384,106
FILING DATE: 06-FEB-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0556 0500PC01
TELEPHONE: 202-371-2500
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
TRANSDUCED: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: p19
NAME/KEY: CDS
LOCATION: 1..501
PCT-US96-01643-8

alignment_scores:
Quality: 74.00 Length: 99
Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303

alignment_block:
US-09-471-276-831 x PCT-US96-01643-8 ..
Align seq 1/1 to: PCT-US96-01643-8 from: 1 to: 501

17 ProvalThrluaAlaIaIle.....PhetygluThrcIn**Serle 31
||||| :||||| ||| :|||||
229 CCATCGATCAGCAGCCGCCGCTGCTGCGACAC..... 287
31 utPalaIuSerGIuHs***LeuYshrlaenGIyInCysaSPAlaA 48
spvalProlGIyPrProGIyAspSerArGIaPro.....AlaVal 61
268 .....CTGAAGTCCTAGCGACACGGGCGTC 295
48 spvalProlGIyPrProGIyAspSerArGIaPro.....AlaVal 61
||||| ||| ||| :|||||
296 ATGCACAGTGCCTGATGGACACGGGCGCTTCATCTCGACGCTT 345
62 GInGlu.....TrpGlyAlaGInGluProva 70
||||| :|||||
346 CAGAGGCGTCACTGCTGTGTCAGCTTTCGACGCTGATATCATCT 395
70 lHsLeuaAspSerProAlaIleYshrlaenGInPheLeuThrGIyasp 86
||||| :||||| ||| :|||||
396 CCATCGACGGAGACCGCGGCGTCTACACCCCTTG..... 430

```

```

87 ThGInClYArGyTArGysArGserGlyIenSeThrGlyTrp 101
431 .....ACCTGGACTCCAGAGAGGAGGCTCCAGACCTCTGG 466
seq_name: /cgm2_6/prodbt/2/lna/6A.COMB.seq:US-08-384-106A-7
seq_documentation_block:
: Sequence 7, Application US/08384106A
: Patent No. 6033847
: GENERAL INFORMATION:
: APPLICANT: Sherr Ph.D., Charles J.
: APPLICANT: Sher Ph.D., James
: APPLICANT: Inokai Ph.D., Hiroshi
: APPLICANT: Okada, Tsukasa
: TITLE OF INVENTION: InK4c-p18 and InK4d-p19, inhibitors of
: NUMBER OF INVENTION: Cyclin Dependent Kinases CDK4 and CDK6, and Uses Thereof
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
: STREET: 1100 NEW YORK AVE., N.Y.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: DOS/PC/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/384,106A
: FILING DATE: 06-FEB-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REGISTRATION NUMBER: 30,353
: REFERENCE/DOCKET NUMBER: 0656, 0500000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2540
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS
: LENGTH: 555 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-384-106A-7

alignment_scores:
: Quality: 74.00 Length: 99
: Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303

alignment_block:
US-09-471-276-831 x US-08-384-106A-7 ..
Align seg 1/1 to: US-08-384-106A-7 from: 1 to: 555

17 ProvalThrlAlaAlaIle.....PhetyGluThrGln***Series 31
||||| :|||||
283 CAGATCGATACGACCGACCGACCGATTCCTGACACG.....321
31 UTRPalAcIaSeGtUnHs***LelUyThrlaGlyVcIaGysArPaIaA 48
322 .....CTGAAGTCTCACTGACGACCGGCGGCTG 349
48 sryAlPrGIGrPrGPrGtIyAaSeGrIyGlePrG.....AlaVal 61
||||| ||| :|||
350 ATGTCAACGTGCTGATCGACCGACCGACCGATTCGAATCGATCTGGACATT 399
62 GInClY.....TPrGlyAlaGInIaPrGva 70

```

```

|||||
400 CAGAGAGGTCACACTGCTGCTGACCTTCTGACCTGAACTGATATC 449
70 LHS-LeaspSerPrGPrGtIyLelUySHlGclnPhleIaGleUthrGlyAsp 86
|||||
450 CACTCGACGACGACCGACCGGCTCTGACACCTTGG.....484
87 ThGInClYArGyTArGysArGserGlyIenSeThrGlyTrp 101
485 .....ACCTGGACTCCAGAGAGGAGGCTCCAGACCTCTGG 520
seq_name: /cgm2_6/prodbt/2/lna/6A.COMB.seq:US-08-384-106A-19
seq_documentation_block:
: Sequence 19, Application US/08384106A
: Patent No. 6033847
: GENERAL INFORMATION:
: APPLICANT: Sherr Ph.D., Charles J.
: APPLICANT: Downing Ph.D., James
: APPLICANT: Inokai Ph.D., Hiroshi
: APPLICANT: Okada, Tsukasa
: TITLE OF INVENTION: InK4c-p18 and InK4d-p19, inhibitors of
: NUMBER OF INVENTION: Cyclin Dependent Kinases CDK4 and CDK6, and Uses Thereof
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
: STREET: 1100 NEW YORK AVE., N.Y.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: DOS/PC/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/384,106A
: FILING DATE: 06-FEB-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REGISTRATION NUMBER: 30,353
: REFERENCE/DOCKET NUMBER: 0656, 0500000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2540
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS
: LENGTH: 553 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 55..552
: US-08-384-106A-19

alignment_scores:
: Quality: 74.00 Length: 99
: Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303

alignment_block:
US-09-471-276-831 x US-08-384-106A-19 ..
Align seg 1/1 to: US-08-384-106A-19 from: 1 to: 555

17 ProvalThrlAlaAlaIle.....PhetyGluThrGln***Series 31
||||| :|||||
283 CAGATCGATACGACCGACCGACCGATTCCTGACACG.....321

```

```

31 utPrAlaGluSerGluHis**LeuLyThrLeuGlyGlnCysAspAlaA 48
322 .....CTGAAGGCTCTCACTGAGACACGGGCTG 349
48 sPvAlProGlyProProGlyAspSerGlyLeuPro.....AlaVal 61
||||| ||||| ||||| ..... ||||| |||||
350 ATGTCACAGCTGCTGATGCGACCGGGCACTTCACATTCACCTGGACCT 399
62 GlnGlu.....TTCGlyAlaGlnGlyProVa 70
400 CAAAGAGGCTCACACTGCTGTGCTGACCTTGTCCACAGTCAATCTGACT 449
||||| ||||| ||||| ..... ||||| |||||
70 LHisLeuAspSerProAlaLLeuLyHisGlnInPheLeuLeuThrGlyAsp 86
||||| ..... ||||| ..... ||||| ..... |||||
seq_name: /cgn2_6/prodata/2/lna/5b_comb.seq-us-09-240-906-1
seq_documentation_block:
: sequence 1, Application: us/09240906
: Patent No. 6245965
: GENERAL INFORMATION:
: APPLICANT: ROUSSEL, MARITNE F.
: APPLICANT: ROUSSEL, MARITNE F.
: APPLICANT: ZINDY, FREDERIQUE
: APPLICANT: CUNNINGHAM, JUSTINE
: TITLE OF INVENTION: CELLS THAT LACK P13K1D AND P27K1P1 ACTIVITY AND
: FILE REFERENCE: 1340-1-025
: CURRENT APPLICATION NUMBER: US/09/240_906
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 1
: LENGTH: 555
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-240-906-1

```

```

Alignment_scores:
Quality: 74.00 Length: 99
Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303
Alignment_block:
US-09-471-276-831 x US-09-240-906-1
Align seg 1/1 to: US-09-240-906-1 from: 1 to: 555

```

```

450 CCATCGACGGGACCCGACGGGCTCTCACACCTTGG..... 484
87 ThrGlnGlyArgGlyArgCysArgSerGlyLeuSerThrGlyTrp 101
485 .....ACGTGGACCTCCAGAGAGGGGCTCCAGACACCTCTCTGG 520
seq_name: /cgn2_6/prodata/2/lna/PCTUS_COMB.seq-PCT-US96-01643-7
seq_documentation_block:
: sequence 7, Application: PC/TUS9601643
: GENERAL INFORMATION:
: APPLICANT: St. Jude Children's Research Hospital
: TITLE OF INVENTION: Ink4c-p18 and Ink4d-p19, inhibitors of
: NUMBER OF SPIROCKS: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
: STREET: 1100 NEW YORK AVE., N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: PC/US9601643
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/01643
: CLASSIFICATION:
: PRIORITY INFORMATION:
: PRIOR APPLICATION DATA:
: FILING DATE: 08/384.106
: APPLICATION NUMBER: 06-FEB-1995
: ATTORNEY/AGENT INFORMATION:
: REGISTRATION NUMBER: 30.353
: REFERENCE/DOCKET NUMBER: 0656.050PC01
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2540
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 555 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: POSITION IN GENOME:
: ORIGIN: cDNA/STROMA/STROMA: p19
PCT-US96-01643-7

```

```

Alignment_scores:
Quality: 74.00 Length: 99
Ratio: 1.542 Gaps: 6
Percent Similarity: 48.485 Percent Identity: 30.303
Alignment_block:
US-09-471-276-831 x PCT-US96-01643-7
Align seg 1/1 to: PCT-US96-01643-7 from: 1 to: 555

```

```

17 ProValThrGlnAlaAlaLe.....PheGlyGlnThrGln***Serie 31
||||| ..... ||||| ..... ||||| ..... |||||
283 CCAGTCTCAGTACAGCCGACCCGACCTGATCTCTGAGACAC..... 321
31 utPrAlaGluSerGluHis**LeuLyThrLeuGlyGlnCysAspAlaA 48
||||| ..... ||||| ..... ||||| ..... |||||
322 .....CTGAAGGCTCTCACTGAGACACGGGCTG 349
48 sPvAlProGlyProProGlyAspSerGlyLeuPro.....AlaVal 61
||||| ||||| ||||| ..... ||||| |||||
350 ATGTCACAGCTGCTGATGCGACCGGGCACTTCACATTCACCTGGACCT 399
62 GlnGlu.....TTCGlyAlaGlnGlyProVa 70
||||| ..... ||||| ..... ||||| ..... |||||
400 CAAAGAGGCTCACACTGCTGTGCTGACCTTGTCCACAGTCAATCTGACT 449
||||| ..... ||||| ..... ||||| ..... |||||
70 LHisLeuAspSerProAlaLLeuLyHisGlnInPheLeuLeuThrGlyAsp 86
||||| ..... ||||| ..... ||||| ..... |||||
322 .....CTGAAGGCTCTCACTGAGACACGGGCTG 349

```

```

48 sptvalprocllyprocllyasptserarglaupro.....AlaVal 61
||||| 111 111 111 111 111 111 111 111 111 111
350 ATCTCACGTCGCTGATGGACCGGACCTCTGCATTCACATCCATCTGCAGATT 399
62 GlnGlu .....TPrGlyAlaGlnGluProva 70
||||| 111 111 111 111 111 111 111 111 111 111
400 CAGAGGCTCACACTGCTGTCACCTTCTGACAGCTGATATCATCT 449
70 1His-LeuAspSerProAla1leIySallIaGlnPheLeuIleuThrlGlyAsp 86
||||| 111 111 111 111 111 111 111 111 111 111
450 CATTCCGACGACACACAGGAGCTCTACACCTTTGG..... 484
87 ThrGlnGlyArgTyrArgCysArgSerGlyLeuSerThrGlyTyr 101
||||| 111 111 111 111 111 111 111 111 111 111
485 .....ACGTGGACTCTCAGAGAGGGGCTCTACGACCTGCCTGG 520

```

```

seq_name: /cgn2_6/prodata/2/1na/PCTUS_COMB_seq:PCT-US96-05252-1

```

seq_documentation block:

```

: Sequence 1 Application: PC/US9605252

```

```

: GENERAL INFORMATION:

```

```

: APPLICANT:

```

```

: TITLE OF INVENTION: p19: A Cell Cycle Inhibitor

```

```

: NUMBER OF SEQUENCES: 6

```

```

: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Townsend and Townsend and Crew

```

```

: STREET: One Montgomery Plaza, Stearns Street Tower

```

```

: CITY: San Francisco

```

```

: STATE: California

```

```

: COUNTRY: USA

```

```

: ZIP: 94105-1492

```

```

: COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk

```

```

: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

: CURRENT APPLICATION DATA:

```

```

: FILING DATE: Not yet assigned

```

```

: CLASSIFICATION:

```

```

: PRIOR APPLICATION DATA:

```

```

: PATENT APPLICATION NO.: US 08/425,093

```

```

: FILING DATE: 17-Apr-1995

```

```

: ATTORNEY/AGENT INFORMATION:

```

```

: NAME: Weber, Ellen L.

```

```

: REGISTRATION NUMBER: 32,762

```

```

: REFERENCE/DOCKET NUMBER: 02307B-059910PC

```

```

: TELEPHONE: (415) 513-9600

```

```

: TELEFAX: (415) 513-9600

```

```

: INFORMATION FOR SEQ ID NO: 1:

```

```

: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 706 base pairs

```

```

: TYPE: nucleic acid

```

```

: STRANDEDNESS: single

```

```

: TOPOLOGY: Linear

```

```

: MAPPING INFORMATION:

```

```

: FEATURE: TYPE: DNA (genomic)

```

```

: NAME/KEY: CDS

```

```

: LOCATION: 1..498

```

```

: OTHER INFORMATION: /product: "human p19"

```

```

: PCT-US96-05252-1

```

```

: Alignment_scores:

```

```

: Quality: 74.00

```

```

: Ratio: 1.542

```

```

: Percent Similarity: 48.485

```

```

: Percent Identity: 30.303

```

```

: Alignment_block:

```

```

: US-09-471-276-831 x PCT-US96-05252-1

```

```

Align seq 1/1 to: PCT-US96-05252-1 from: 1 to: 706
17 ProvalThrGluAlaIalle.....PheTyrGluThrGln***SerLe 31
||||| 111 111 111 111 111 111 111 111 111 111
229 CATTCCATCAGACACACCGGACCTCTGCATTCCTGCACAC..... 267
31 uTPAlaGluSerGluHis**LeuYsthrLeuAlYgInCysAspAlaA 48
||||| 111 111 111 111 111 111 111 111 111 111
268 .....CTGAGGCTCTCAGACACGCGAGCTT 295
48 sptvalprocllyprocllyasptserarglaupro.....AlaVal 61
||||| 111 111 111 111 111 111 111 111 111 111
296 ATCTCACGTCGCTGATGGACCGGACCTCTGCATTCACATCCATCTGCAGATT 345
62 GlnGlu .....TPrGlyAlaGlnGluProva 70
||||| 111 111 111 111 111 111 111 111 111 111
346 CAGAGGCTCACACTGCTGTCACCTTCTGACAGCTGATATCATCT 395
70 1His-LeuAspSerProAla1leIySallIaGlnPheLeuIleuThrlGlyAsp 86
||||| 111 111 111 111 111 111 111 111 111 111
396 CATTCCGACGACACACAGGAGCTCTACACCTTTGG..... 430
87 ThrGlnGlyArgTyrArgCysArgSerGlyLeuSerThrGlyTyr 101
||||| 111 111 111 111 111 111 111 111 111 111
431 .....ACGTGGACTCTCAGAGAGGGGCTCTACGACCTGCCTGG 466

```

```

seq_name: /cgn2_6/prodata/2/1na/76_COMB_seq:US-08-667-939A-13

```

seq_documentation block:

```

: Sequence 13 Application: US/08667939A

```

```

: Patent No. 5998166

```

```

: GENERAL INFORMATION:

```

```

: APPLICANT: LHO, Shun

```

```

: TITLE OF INVENTION: CD16-II VARIANTS

```

```

: NUMBER OF SEQUENCES: 5

```

```

: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: BROWDY AND NEIMARK

```

```

: STREET: 419 Seventh Street, N.W., Suite 300

```

```

: CITY: Washington

```

```

: STATE: D.C.

```

```

: COUNTRY: USA

```

```

: ZIP: 20004

```

```

: COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk

```

```

: OPERATING SYSTEM: IBM PC compatible

```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

: CURRENT APPLICATION DATA:

```

```

: FILING DATE: 24-JUN-1996

```

```

: PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER: US 08/433,123

```

```

: FILING DATE: 03-MAY-1995

```

```

: ATTORNEY/AGENT INFORMATION:

```

```

: NAME: BROWDY, Roger L.

```

```

: REGISTRATION NUMBER: 25,618

```

```

: REFERENCE/DOCKET NUMBER: LHO-2A

```

```

: TELEPHONE: 202-628-5197

```

```

: TELEFAX: 202-737-3528

```

```

: INFORMATION FOR SEQ ID NO: 13:

```

```

: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 765 base pairs

```

```

: TYPE: nucleic acid

```

```

: STRANDEDNESS: single

```

```

: TOPOLOGY: Linear

```

```

: MOLECULE TYPE: cDNA

```

```

: US-08-667-939A-13

```

```

: Alignment_scores:

```

```

: Quality: 74.00

```

```

: Ratio: 1.423

```

```

: Length: 91

```

```

: Gaps: 7

```

Percent Similarity: 57.143 Percent Identity: 35.165

Alignment_block:

US-09-471-276-831 x US-08-667-939A-13 ..

Align seg 1/1 to: US-08-667-939A-13 from: 1 to: 765

```

43 G1G1nCysAspAlaAspValProCly.....ProProGly Asps 56
|||||.....|||||.....||| ||| |||
120 GAGAGGTGTACTGTGAGAGGCCAGGGAGCCACTACCTCCCGGAGAGCAATT 169
56 efaThgLeuProAlaValaGlnGluTrrpGlyAlaGlnGluProValaHisLeu 72
|||||.....|||||.....|||||
170 CC.....ACGAGATGCTTTCACAAAGC.....ACCTC 198
73 AspSerProAlaIleuYahIleGlnPhela.....LeuThGlyAspTh 87
|||||.....|||||.....|||||
199 ATCTCAAGCCAGGCGCTGAGCTACTTATGACCTGCCAGCAGTCAGCA 248
87 rGlnGlyATrGATrGcYATrGcATrGcLeuSerThrGlyTrp**GlnL 104
|||||.....|||||.....|||||
249 CAGTGGAGACTACAGTCCAGGCGAGCAAGCTCTCCAGC.....C 286
104 euSerLyLeuLeuGlnLeu..ThrGlyProValaLeuAlaCysSerLe 120
|||||.....|||||.....|||||
287 TCAGTACCCGCGGTGACGCTGAGAGATCCAGTCGCGCTGCTGTCTCCAG 336
120 uAlaLeuAspGlyAlaSer 126
|||||.....|||||
337 GCCCTCGTGGGTCTTCA 355

```

seq_name: /cgn2_6/prodata/2/lna/5E_COMB.seq:US-08-667-939A-14

seq_documentation_block:

```

: Sequence No. 595330
: GENERAL INFORMATION:
: APPLICANT: LHO, Shun
: TITLE OF INVENTION: CD16-II VARIANTS
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEWMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: STATE: Washington
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: SYSDAT in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/667,939A
: FILING DATE: 24-JUN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/433,123
: FILING DATE: 03-MAY-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY AND NEWMARK
: REGISTRATION NUMBER: 55,618
: REFERENCE/DOCKET NUMBER: LHO-2A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 286 bp
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-667-939A-14

```

alignment_scores:

Quality: 74.00 Length: 91

Percent Similarity: 57.143 Percent Identity: 35.165

Alignment_block:

US-09-471-276-831 x US-08-667-939A-14 ..

Align seg 1/1 to: US-08-667-939A-14 from: 1 to: 765

```

43 G1G1nCysAspAlaAspValProCly.....ProProGly Asps 56
|||||.....|||||.....||| ||| |||
120 GAGAGGTGTACTGTGAGAGGCCAGGGAGCCACTACCTCCCGGAGAGCAATT 169
56 efaThgLeuProAlaValaGlnGluTrrpGlyAlaGlnGluProValaHisLeu 72
|||||.....|||||.....|||||
170 CC.....ACGAGATGCTTTCACAAAGC.....ACCTC 198
73 AspSerProAlaIleuYahIleGlnPhela.....LeuThGlyAspTh 87
|||||.....|||||.....|||||
199 ATCTCAAGCCAGGCGCTGAGCTACTTATGACCTGCCAGCAGTCAGCA 248
87 rGlnGlyATrGATrGcYATrGcATrGcLeuSerThrGlyTrp**GlnL 104
|||||.....|||||.....|||||
249 CAGTGGAGACTACAGTCCAGGCGAGCAAGCTCTCCAGC.....C 286
104 euSerLyLeuLeuGlnLeu..ThrGlyProValaLeuAlaCysSerLe 120
|||||.....|||||.....|||||
287 TCAGTACCCGCGGTGACGCTGAGAGATCCAGTCGCGCTGCTGTCTCCAG 336
120 uAlaLeuAspGlyAlaSer 126
|||||.....|||||
337 GCCCTCGTGGGTCTTCA 355

```

seq_name: /cgn2_6/prodata/2/lna/5E_COMB.seq:US-08-483-376-1

seq_documentation_block:

```

: Sequence No. 595330
: Patent No. 595330
: GENERAL INFORMATION:
: APPLICANT: Vasili, Yulia
: APPLICANT: Clancy, Maureen A.
: APPLICANT: Vasil, Robert Y.
: APPLICANT: Vasil, Robert Y.
: APPLICANT: Hannah, L.C.
: TITLE OF INVENTION: No. 595330el Means for Enhancing Gene
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: Boulder
: CITY: Boulder
: STATE: Colorado
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: SYSDAT in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/483,376
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/418,540
: FILING DATE: 04-AUG-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/102,115
: FILING DATE: 04-AUG-1993
: PRIOR APPLICATION DATA:

```

```

alignment_scores:
    quality: 73.00          Length: 149
    Ratio: 1.109           Gaps: 8
    Percent Similarity: 44.295   Percent Identity: 20.805

Alignment_block:
US-09-471-276-83j x US-06-483-376-J/rev ..

Align seg 1/1 to reverse of: US-06-483-376-1 from: 1 to: 6386

      |||
      |||
11 TgcTgATGThTTTCgtCTGTGCtTAAThAGAlAAllethyrtctm 27
|||
5643 tgg...cttcgccgatattccGAAGAAGAcCGCtCArTGATTcACAC 5597
27 tcGa**serLeurPtlagIsorGtJshH**Lau..... 39
5596 cOTTAAACtACTGGAAATGGAGAACAGACtCATGAtCTCTCTCACAGCC 5547
|||||
39 ..... 39
5546 AAGGGAGAGGACGAGGCCAGCACCATAAGCGCTTACAAGCGTAGC 5497
40 .....LysThrLeuGlYglncysAspaIalaspALPrpolYPProgrGI 54
|||||
5496 TCAGGCGCTGGAACACTGTGATGTAGCGCGCGCTGTGCGCCTCCACAG 5447
54 yAgSPeRqR lePR..AlayAdnIdntRPclYAldnglunpg 69
|||||
5446 TTGCTCACTACTCTCGAAGACCGCAACGCCCGTCAAGGTCACTACGT 5397
70 ValHislnslaspsEPNoIAlelyshlsGLIn..... 80
5396 CTCGGAGTAAGAGCTGTACGTGACTGACGACATCAAACCTGTAGCT 5347
81 .....PhdEnteurhtyAaphrhtnlngLVArGrVAr 92
5346 CAGATTTTTTACACTCTCTCTTCTTCTACTAGAGAAAGCGAGATTC 5297
92 rGcYatSerGcyJleusErThGtYTrP**GJleuseSrtyLeoudeu 108
|||||
5296 CGAGCTCCGCTGGAGTAACAGTCCCTCTGATCAACAATAATTT 5247
109 GlueUenrhGtPrOxySethlaIdnaICyseSeSTheAlaleuasP 123
|||||
5246 ACATTCTGACGAGGAGAAAAAAAATATCATCTACTTCACVTAAr 5202

seq_name: /cgcn2_6/prodaaa/2/lua/6A_cowB.seq.us-09-193-191-3

seq_documentation_block:
# Sequences used in alignment
# Patient No. 6130078
# GENERAL INFORMATION:
# APPLICANT: XIAO, JIAN-PING
# APPLICANT: XU, SHUAN-YONG
# TITLE OF INVENTION: METHOD FOR CLONING THE NSPRT RESTRICTION-MODIFICATION SITE RESPONSIBLE FOR PRODUCING THE RECOMBINANT NSPHT
# TITLE OF INVENTION: SYSTEM IN E. COLI AND PRODUCING THE RECOMBINANT NSPHT
# CURRENT APPLICATION NUMBER: US/09/193,191
# NUMBER OF SEQ ID NOS: 12
# SOFTWARE: PatentIn Ver. 2.0
# SEQ ID NO 3
# LENGTH: 735
# TYPE: DNA
# SOURCE: GenBank
# ACCESSION: No. G130078LOC SP.
# FEATURE:
# NAME/KEY: CDS
# LOCATION: (1)..(735)
# US-09-193-191-3
```


Alignment scores:
 Quality: 72.50 Length: 56
 Ratio: 1.959 Gaps: 4
 Percent Similarity: 66.071 Percent Identity: 37.500

alignmentL_block:
 US-09-471-276-831 x US-09-193-191-3 ..

Align seq 1/1 to: US-09-193-191-3 from: 1 to: 735

```

55 ASPPSRATGLEPRLAVALGAGLGLITTPGVALAGLGLPRLVLAH1 71
   |||||.....|.....|.....|.....|.....|.....|
394 GATCATCATCTCCG...ACGACAGCTGGAGGCTTACAGACCTCCCA 440
   |||||...|...|...|...|...|...|...|...|...|...|
71 SLAASPSRPRPLA...||ELYAHISGLPHE...LEULAUT 84
   |||||...|...|...|...|...|...|...|...|...|...|
441 CTTAACCTCTGATGAGATGAATAATTACCAAAAGTTTCATGTTTAA 490
   |||||...|...|...|...|...|...|...|...|...|...|
84 hrclyasprhinclyatgtytarg.....Cysarg 94
   |||||...|...|...|...|...|...|...|...|...|...|
491 AAAAAAGCATACAGGAGTACAGATCTTTTAAATGCGAGAGCTGGAG 540
   |||||...|...|...|...|...|...|...|...|...|...|
95 SerGlyLeuSerThrGly 100
   |||||...|...|...|...|...|...|...|...|...|...|
541 CCCTGATCAAAAGCTGT 558

```

seq_name: /cgn2_6/prodata/2/lna/5A.COMB.seq:US-07-668-648-3

seq documentation block:
 ? Sequence 3, Application US/07668648
 ? Patent No. 5416192
 ? GENERAL INFORMATION:
 ? APPLICANT: Shoyab, Mohammed
 ? APPLICANT: Plovman, Gregory D.
 ? TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH
 ? NUMBER OF INVENTORS: 2
 ? CORRESPONDENCE ADDRESS: 112
 ? ADDRESSEE: Pennle & Edmonds
 ? STREET: 1155 Avenue of the Americas
 ? CITY: New York
 ? STATE: New York
 ? COUNTRY: USA
 ? PRIORITY: 003
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? OPERATING SYSTEM: IBM PC COMPATIBLE
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25
 ? CURRENT APPLICATION DATA: 648
 ? APPLICATION NUMBER: US/07/668-648
 ? PUBLICATION NUMBER: 591819
 ? CLASSIFICATION: 514
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Mistrock, S. Leslie
 ? REGISTRATION NUMBER: 18,872
 ? REFERENCE/DOCKET NUMBER: 5624-161-999
 ? TELEPHONE: (212) 900-9090
 ? TELEFAX: (212) 650-9090
 ? INFORMATION FOR SEQ ID NO: 3:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1779 base pairs
 ? TYPE: NUCLEIC ACID
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MODIFICATION: none
 ? ORIGIN: SOURCE: cDNA
 ? ORGANISM: Homo sapiens
 ? TISSUE TYPE: Kidney
 ? FEATURE:
 ? NAME/KEY: CDS

LOCATION: 1..1779
 US-07-668-648-3

Alignment scores:
 Quality: 71.50 Length: 140
 Ratio: 1.172 Gaps: 6
 Percent Similarity: 43.571 Percent Identity: 23.571

alignmentL_block:
 US-09-471-276-831 x US-07-668-648-3 ..

Align seq 1/1 to: US-07-668-648-3 from: 1 to: 1779

```

10 LEUTPGVALVLAHTRP.....GLYPRVALTRPVALAHISLAAPS 23
   |||||.....|...|...|...|...|...|...|...|...|...|
80 CTGGGCGCTCTGCGTGGAGCCGAGAGAGAGACAGTACAGCTGCGCT 129
   |||||...|...|...|...|...|...|...|...|...|...|
23 EPIETRYLTHRGH.....**SerLeuTrpAlaGluS 35
   |||||...|...|...|...|...|...|...|...|...|...|
130 CCCTCTGCGCAATAGCCGACCCGACAGACAGACAGATGCTGGATGCC 179
   |||||...|...|...|...|...|...|...|...|...|...|
35 ERGLHIS**LeuIysThr...LeuGlyGlnCysAspAlaAspValPro 50
   |||||...|...|...|...|...|...|...|...|...|...|
180 CTGGCAGGTGATGCCACTGCTCTGCGCGGAGCCGACATCTGATCTAC 229
   |||||...|...|...|...|...|...|...|...|...|...|
51 GLYPRPGLYASPRATGLEPRLAVALGAGLGLITTPGVALAGLGLP 64
   |||||...|...|...|...|...|...|...|...|...|...|
230 TCTGACGAGATTCAGATGTTGCTGCCCTTCCAGAGGCGCTGACATGG 279
   |||||...|...|...|...|...|...|...|...|...|...|
65 .....GLYALAGLGLPRLVLAHISLAAPS 74
   |||||...|...|...|...|...|...|...|...|...|...|
280 GATGCGCATCACTGCTGCCGACGAGGCTTCCAGATGTCGACAGCGCG 329
   |||||...|...|...|...|...|...|...|...|...|...|
74 ETPRLAALILEYSHISGLINPHELEULEUTHRGLYASPRHINCGLYAT 90
   |||||...|...|...|...|...|...|...|...|...|...|
330 ATGCTGCTCTCAAAAGTACG.....349
   |||||...|...|...|...|...|...|...|...|...|...|
91 TYRAGCYSARGSERGLYLEUSERTHRGTYRTP**GlnLeuSerIysLe 107
   |||||...|...|...|...|...|...|...|...|...|...|
350 .....GTACACACTGCGTGGGTGCATCGACTGCTCT 381
   |||||...|...|...|...|...|...|...|...|...|...|
107 ULGU.....GlnLeuThrGlyProLyValLeuAlaCysSerTrp 120
   |||||...|...|...|...|...|...|...|...|...|...|
382 GATGTCAGATTCGATGATGCCGACGAGCTTCCAGAGGCTGCTGATG 431
   |||||...|...|...|...|...|...|...|...|...|...|
120 EULALEUASPCGLYALASER 136
   |||||...|...|...|...|...|...|...|...|...|...|
432 TGGCTCTCGGGGCTGCGCC 451

```

seq_name: /cgn2_6/prodata/2/lna/5A.COMB.seq:US-08-429-998-3

seq documentation block:
 ? Sequence 3, Application US/08429998
 ? Patent No. 5885961
 ? GENERAL INFORMATION:
 ? APPLICANT: Shoyab, Mohammed
 ? APPLICANT: Plovman, Gregory D.
 ? TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH
 ? NUMBER OF INVENTORS: 2
 ? CORRESPONDENCE ADDRESS: 112
 ? ADDRESSEE: Pennle & Edmonds
 ? STREET: 1155 Avenue of the Americas
 ? CITY: New York
 ? STATE: New York
 ? COUNTRY: USA
 ? PRIORITY: 003
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? OPERATING SYSTEM: IBM PC COMPATIBLE
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25

```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/429,998
? FILING DATE: 27-APR-1995
? PRIORITY DATE: 27-APR-1995
? PRIORITY NUMBER: 276-831
? APPLICATION NUMBER: US 07/668,648
? FILING DATE: 13-MAR-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Mistrock, S. Leslie
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 5624-161-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-9741
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1779 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? TISSUE TYPE: kidney
? FEATURE:
? NAME/KEY: CDS
? ORIGIN: 1..1779
US-08-429-998-3

```

```

alignment_scores:
? Quality: 71.50 Length: 140
? Ratio: 1.172 Gaps: 6
? Percent Similarity: 43.571 Percent Identity: 23.571

```

```

alignment_block:
US-09-471-276-831 x US-08-429-998-3 ..

```

```

Align seg 1/1 to: US-08-429-998-3 from: 1 to: 1779
10 LeuTgclgylvalthtrp.....glyprvvalthgclvalalal 23
80 CTCTGCGCTCTGCGCGCGCGAGAGCGAGCGCGCGCGCGCGCGCG 129
23 ephetrpvclythrcln.....**serleuprpalagus 35
35 egrclnlns**leuclgthr..leuclgclncysaapalaaprrp 50
180 CTGCGCGCTGTAATGCGAGCTCTGCGCGCGCGCGCGCGCGCGCG 229
51 glprprrpdluaarsafragluarpalavaglnclnlytrp..... 64
230 TCTCAGCGAGCTTCCAGCTGCGCGCTCTGCGCGAGCGCGCGCGCG 279
65 .....gylasglncluprvvalhllsleuasp 74
280 GATGCGCGAGCTGCGCGCGCGAGCGCGCTCTGCGCGCGCGCGCG 329
74 egrpvalalaleuvalnaglnrphyleuclvthrcldyaarhclnclyar 90
330 ATCTCTGCTTCCAAAGTCA.....**serleuprpalagus 349
91 tyaarpruaarvserclyleuclgthr**cgluclserclvle 107
350 .....gtaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 381
107 uleu .....gtaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 381
382 GATGCGCGAGCTGCGCGCGCGAGCTCTGCGCGCGCGCGCGCGCG 431
120 gualaleuarpdlualaser 126

```

```

seq.name: /cgn2_6/plodata/2/lna/56.COMB.seq:US-08-431-333-3
432 TGCGTCTGCGGCGTCTGCC 431
seq documentation block:
? Sequence 3, Application US/08431333
? Patent No. 5965723
? GENERAL INFORMATION:
? APPLICANT: Shoyman, Mohammed
? APPLICANT: Ploymann, Gregory D.
? TITLE OF INVENTION: BETHELINS: NOVEL CYSTEINE-RICH GROWTH
? NUMBER OF INVENTION: BETHELINS: NOVEL CYSTEINE-RICH GROWTH
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Penelle & Edmonds
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10036
COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
? FILING DATE: 27-APR-1995
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/668,648
? FILING DATE: 13-MAR-1991
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 5624-161-999
TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-9741
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1779 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? TISSUE TYPE: kidney
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1779
US-08-431-333-3

```

```

alignment_scores:
? Quality: 71.50 Length: 140
? Ratio: 1.172 Gaps: 6
? Percent Similarity: 43.571 Percent Identity: 23.571

```

```

alignment_block:
US-09-471-276-831 x US-08-431-333-3 ..

```

```

Align seg 1/1 to: US-08-431-333-3 from: 1 to: 1779
10 LeuTgclgylvalthtrp.....glyprvvalthgclvalalal 23
80 CTCTGCGCTCTGCGCGCGAGAGCGAGCGCGCGCGCGCGCGCG 129
23 ephetrpvclythrcln.....**serleuprpalagus 35
35 egrclnlns**leuclgthr..leuclgclncysaapalaaprrp 50
180 CTGCGCGCTGTAATGCGAGCTCTGCGCGCGCGCGCGCGCGCG 229
51 glprprrpdluaarsafragluarpalavaglnclnlytrp..... 64
230 TCTCAGCGAGCTTCCAGCTGCGCGCTCTGCGCGAGCGCGCGCGCG 279
65 .....gylasglncluprvvalhllsleuasp 74
280 GATGCGCGAGCTGCGCGCGCGAGCGCGCTCTGCGCGCGCGCGCG 329
74 egrpvalalaleuvalnaglnrphyleuclvthrcldyaarhclnclyar 90
330 ATCTCTGCTTCCAAAGTCA.....**serleuprpalagus 349
91 tyaarpruaarvserclyleuclgthr**cgluclserclvle 107
350 .....gtaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 381
107 uleu .....gtaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 381
382 GATGCGCGAGCTGCGCGCGCGAGCTCTGCGCGCGCGCGCGCGCG 431
120 gualaleuarpdlualaser 126

```



```

? APPLICATION NUMBER: US/08/464,517
? FILING DATE:
? CLASSIFICATION: 435
? PRIORITY NUMBER: 07/963,308
? PRIOR APPLICATION DATA: US 07/963,308
? FILING DATE: 16-OCT-1992
? APPLICATION NUMBER: US 07/888,178
? FILING DATE: 26-MAY-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/701,514
? FILING DATE: 16-MAY-1991
? ATTORNEY FIRM: J. DOMANUSKI
? NAME: Matthew P. Vincent
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: MIT-004C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 35:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1624 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-464-517-35

```

```

alignment_scores:
? Quality: 71.00 Length: 101
? Ratio: 1.315 Gaps: 4
? Percent Similarity: 53.465 Percent Identity: 26.733

```

alignment_block:

US-09-471-276-831 x US-08-464-517-35/rev ..

Align seg 1/1 to reverse of: US-08-464-517-35 from: 1 to: 1624

```

8 LeuLeuLeuTrp.GlyValThrTrpGlyProValThrGluAlaAlaLeu 24
   :::::::::::::::::::: ||| ::::::::::::::
674 GTGCTTCGTGAGAGGATGATGATGATGATGATGATGATGATGATGCTGT 625
24 hefyrcilthrgln***seletrpAlaIglusrciluhls***leuLys 40
   :::::::::::::::::::: ||| ::::::::::::::
624 TTTCGGGGGCTCTTCGGGGGCTTCGGGGGCGAGCCG.....GTCCAG 584
41 ThrLeuGlyGlnCysAspAlaAspAlaProGlyProProGlyAspSerLeu 57
   :::::::::::::::::::: ||| ::::::::::::::
583 CCGTCTAGAGCCCGCCAGCGCGAGCGAGCGCTCCCGACCGCTCCCGCC 534
57 glenProAlaIvalGlnGlnTrpGlyAlaGlnGlnProValHisLeu... 72
   :::::::::::::::::::: ||| ::::::::::::::
533 GCACGCCGCGCAGCGAGGAAACCGCTTCCTCCGCGCTTCGATCTGTGCA 484
73..AspSerPro.....AlaIleuYnHisGlnPheLeuLeu 83
   :::::::::::::::::::: ||| ::::::::::::::
483 CACGCCCGCCGAGGTCTCTCGCTGACGCTGCGACCGACCGACGATGACGCTTT 434
84 ThrLeuAspThrGlnGlnIleuTrpGlyAlaGlnGlnSerGlyLeuSerThrGln 100
   :::::::::::::::::::: ||| ::::::::::::::
433 CTGCGAAATGCCCCCGAGCGAGAGCGCGAGCGCGAGCGAGCGATTAAGTCTCCG 384
100 Y 100
383 C 383

```

```

seq_name: /cgm2_6/ptcdeta/2/lna/5b_COMB.seq:us-08-246-361A-35
seq_documentation_block:
? Sequence 35, Application US/08246361A
? Patent No. 5998582
? GENERAL INFORMATION:
? APPLICANT: BEACH, David H.

```

```

? TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSER: LAHVE & COCKFIELD
? CITY: 60 State Street
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? SOFTWARE: SYSTEM: DOS/MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/246,361A
? FILING DATE: 19-MAY-1994
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA: US 07/963,308
? FILING DATE: 16-OCT-1992
? APPLICATION DATA:
? APPLICATION NUMBER: US 07/888,178
? FILING DATE: 26-MAY-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/701,514
? FILING DATE: 16-MAY-1991
? ATTORNEY FIRM: J. DOMANUSKI
? NAME: Matthew P. Vincent
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: MIT-004C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 35:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1624 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-246-361A-35

```

```

alignment_scores:
? Quality: 71.00 Length: 101
? Ratio: 1.315 Gaps: 4
? Percent Similarity: 53.465 Percent Identity: 26.733

```

alignment_block:

US-09-471-276-831 x US-08-246-361A-35/rev ..

Align seg 1/1 to reverse of: US-08-246-361A-35 from: 1 to: 1624

```

8 LeuLeuLeuTrp.GlyValThrTrpGlyProValThrGluAlaAlaLeu 24
   :::::::::::::::::::: ||| ::::::::::::::
674 GTGCTTCGTGAGAGGATGATGATGATGATGATGATGATGATGATGCTGT 625
24 hefyrcilthrgln***seletrpAlaIglusrciluhls***leuLys 40
   :::::::::::::::::::: ||| ::::::::::::::
624 TTTCGGGGGCTCTTCGGGGGCTTCGGGGGCGAGCCG.....GTCCAG 584
41 ThrLeuGlyGlnCysAspAlaAspAlaProGlyProProGlyAspSerLeu 57
   :::::::::::::::::::: ||| ::::::::::::::
583 CCGTCTAGAGCCCGCCAGCGCGAGCGCTCCCGACCGCTTCGCGCCCGCC 534
57 glenProAlaIvalGlnGlnTrpGlyAlaGlnGlnProValHisLeu... 72
   :::::::::::::::::::: ||| ::::::::::::::
533 GCACGCCGCGCAGCGAGGAAACCGCTTCCTCCGCGCTTCGATCTGTGCA 484
73..AspSerPro.....AlaIleuYnHisGlnPheLeuLeu 83
   :::::::::::::::::::: ||| ::::::::::::::
483 CACGCCCGCCGAGGTCTCTCGCTGACGCTGCGACCGACCGACGATGACGCTTT 434

```

```

84 ThnGlyaspThnglnGlnYAcgYArserGlyLeuSerThngl 100
      |||::: ||| :::: |||::: |||
433 CTAGGAAATGACCGGAGGAGGAGGCGGATGAGGATTAAGTCCGG 384
100 Y 100
383 C 383

seq_name: /cgn2_6/prodata/2/lna/6A.COMB.seq:US-08-463-772-35

seq_documentation_block:
? Sequence 35, Application US/08463772
? PC/US9305000
? GENERAL INFORMATION:
? APPLICANT: BEACH, David H.
? TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: LAHIVE & COCKFIELD
? STREET: 60 State Street
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: PC-DOS/MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08463772
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/963,308
? FILING DATE: 26-MAY-1992
? APPLICATION NUMBER: US 07/888,178
? FILING DATE: 26-MAY-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/701,514
? FILING DATE: 16-MAY-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: RACHLEW, Vincent
? REGISTRATION NUMBER: 109
? REFERENCE/DOCKET NUMBER: MII-004C
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 35:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1624 base pairs
? TYPE: double
? STRANDEDNESS: double
? TOPOLOGY: Linear
? MOLECULE TYPE: DNA (genomic)
? US-08-463-772-35

alignment_scores:
? Quality: 71.00 Length: 101
? Ratio: 1.315 Gaps: 4
? Percent Similarity: 53.465 Percent Identity: 26.733

alignment_block:
US-09-471-276-831 x US-08-463-772-35/rev. .
Align seq 1/1 to reverse of: US-08-463-772-35 from: 1 to: 1624
8 LeuLeuLeuTrp-GlyValThnTrpGlyProValThnGlnAlaAlaLeuP 24
:::||||||| ||| ||| ::|||::: |||
674 GTGCTTCTCTGAGGAGTAAGATGATCTCCGACGAGAACCATCATGTGT 625

```

```

24 beryGluThrGln***SerLeuTrpAlaGluSerGluHis**LeuIys 40
      || :::: |||::: |||::: |||
624 TTCTGGGGGACCTCTTGGGGCCCGCTGGGGGAGCGCGC.....GTGCAG 584
41 ThnLeuGlnGlnCysAspAlaAspValProGlyProGlyAspSerAr 57
      |||::: |||::: |||::: |||
583 GTCTAGGCGCCGACGAGGAGCGCGGAGGAGGCTTCGACGCTTCGCCGG 534
57 gleuProAlaValGlnGlnTrpGlyAlaGlnGluProValHisLeu.... 72
      |||::: |||::: |||::: |||
533 GCCACGGCGCACGACGAGAAACCCGCTTCTGCGCCCTGCATCTCTCGA 484
73 . AspSerPro.....AlaIleuYshIscIhpeLeuLeu 83
      |||::: |||::: |||::: |||
483 CAGGCCCGCCGAGGAGGTGTCTGCGCTGAGCGTGGCAGACGATCGACGCTT 434
84 ThnGlyaspThnglnGlnYAcgYArserGlyLeuSerThngl 100
      |||::: ||| :::: |||::: |||
433 CTAGGAAATGACCGGAGGAGGAGGCGGATGAGGATTAAGTCCGG 384
100 Y 100
383 C 383

seq_name: /cgn2_6/prodata/2/lna/PCTUS.COMB.seq:PCT-US93-05000-35

seq_documentation_block:
? Sequence 35, Application PC/US9305000
? PC/US9305000
? GENERAL INFORMATION:
? APPLICANT: MITOX
? TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
? NUMBER OF SEQUENCES: 42
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
? STREET: Two Millitia Drive
? CITY: Lexington
? STATE: Massachusetts
? COUNTRY: US
? ZIP: 02173
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: PC-DOS/MS-DOS
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US93/05000
? FILING DATE: 19930525
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/07/888,178
? FILING DATE: 26-MAY-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: RACHLEW, Vincent
? REGISTRATION NUMBER: 32,227
? REFERENCE/DOCKET NUMBER: CSHL01-02A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-861-6240
? TELEFAX: 616-861-9340
? INFORMATION FOR SEQ ID NO: 35:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1624 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: double
? TOPOLOGY: Linear
? MOLECULE TYPE: DNA (genomic)
? PCT-US93-05000-35

alignment_scores:
? Quality: 71.00 Length: 101
? Ratio: 1.315 Gaps: 4
? Percent Similarity: 53.465 Percent Identity: 26.733

```

```

alignment_block:
US-09-471-276-831 x PCT-US93-05000-35/rev ..
Align seg 1/1 to reverse of: PCT-US93-05000-35 from: 1 to: 1624

8 LeuLeuLeuTrp.glyValThrProGlyProValThrGluAlaValIleP 24
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
674 GTCTCTCTCGTAGGGATGATGATGTCCTCCACAGAGAAACATGTGCT 625
24 hGtTcTlurThrLn**SerLeuTrpAlaIleuSerLumIls**LeuLys 40
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
624 TTCTGGGGGCTTTCGGCCCTGGGGGCGAGCGA.....GTGGAG 584
41 ThrLeuGlyLysAlaIleuAspValProGlyProGlyAspSerAla 57
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
533 GCGACGGCGCGAGCGAGGTCTCGCTCGCCGCGACGCGATGACGCTT 484
73 ..AspSerPro.....AlaIleuLysHisGlnPheLeuLeu 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
483 CAGGCGCGCGAGGTCTCGCTCGCCGCGACGCGATGACGCTT 434
84 ThrGlyAspThrGlnGlyThrGlyThrGlyAspSerGlyLysSerThrG 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
433 CTGCGAATGGCGCGCGAGCGAGCGCGCGACGCGATGACGCTT 384
100 Y 100
383 C 383

seq_name: /cgn2_6/prodata/2/line/5B_COMB.seq:us-08-786-527A-1
seq_documentation_block:
: Sequence 1, Application US/08786527A
: Patient No. 5969210
: GENERAL INFORMATION:
: APPLICANT: Sharma and Montanuy
: TITLE OF INVENTION: No. 5969210el Methods for the Characterization of Compounds w/
: NUMBER OF SEQUENCES: 4
: ADDRESSER: MOCHROFF & ADLER, P.C.
: STREET: 8011 Candle Lane
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77071
: COMPUTER RESPONSIBLE ROW:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: Macintosh
: SOFTWARE: Microsoft word for Macintosh
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/786,527A
: FILING DATE: January 21, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Adler, Benjamin A.
: REGISTRATION NUMBER: 35,423
: REFERENCE/DOCKET NUMBER: D5848
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 713-777-2321
: TELEFAX: 713-777-6908
: INVENTION SOURCE: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 403 bp
: TYPE: nucleic acid
: STRANDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE:
: DESCRIPTION: other nucleic acid

```

```

: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: STRAIN:
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: TISSUE TYPE:
: CELL TYPE:
: CELL LINE:
US-08-786-527A-1

alignment_block:
Quality: 70.50 Length: 77
Ratio: 1.533 Gaps: 4
Percent Similarity: 59.740 Percent Identity: 32.468

Alignment_block:
US-09-471-276-831 x US-08-786-527A-1 ..
Align seg 1/1 to: US-08-786-527A-1 from: 1 to: 403

50 ProGlyProGlyAspSerArg.....LeuProAlaIle 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 CCGCGCCCTCTTTAGAGCGCGCGACGCGACCATGCTGCGCCCGCTG 190
61 IGlnGlnTrpGlyAlaGlnGlnGlnProVal.HisLeuAsp.SerProAlaI 77
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 AACCAAGTGAGGTGCTCGACAGCGCTATGGCAGCGCGACGCGCTTCT 240
77 eLysHisGlnPheLeuLeuThrGlyAspThrGlnGlyThrGlyTrpGly 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 CTCTGCGCGACGCTGTGCTTCCCGGAGAGCATGAGAACTGTCMAAGC 290
94 rGSeGlyLysSerThrGlyTTrp**GlnLeuSerLysLeuGlnGln 110
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 GATTCGGGTGGCGCTG.....AdoTCTGTAGCTG 322
111 ThrGlyProLysValLeuAlaLysSer 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
323 CCAAGCGCTTAAGCGCTGCGCTGTGAC 349

seq_name: /cgn2_6/prodata/2/line/5B_COMB.seq:us-08-667-939A-19
seq_documentation_block:
: Sequence 19, Application US/08667939A
: Patient No. 5958166
: GENERAL INFORMATION:
: APPLICANT: LDO, Shun
: TITLE OF INVENTION: CD16-11 VARIANTS
: NUMBER OF SEQUENCES: 53
: ADDRESSER: BROADWAY AND NEWARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER RESPONSIBLE ROW:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/667,939A
: FILING DATE: 22-MAY-1996
: PRIORITY INFORMATION:
: APPLICATION NUMBER: US 08/433,123
: FILING DATE: 03-MAY-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWN, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: LDO-2A

```

```
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-628-5197
? TELEFAX: 202-737-3528
? INFORMATION FOR SEQ ID NO: 19:
? STRAND: 5' to 3'
? LENGTH: 630 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE: CDS
? LOCATION: 7..635
? US-08-667-939A-19

alignment_scores:
? Quality: 70.50 Length: 79
? Ratio: 1.808 Gaps: 6
? Percent Similarity: 49.367 Percent Identity: 34.177

alignment_block:
? US-09-471-276-831 x US-08-667-939A-19 ..
?
? Align seg 1/1 to: US-08-667-939A-19 from: 1 to: 630
?
? 43 GIVGIVCysAspAlaAspValProclY.....ProProclY Asps 56
? 126 GGNACATGTGACTCTGAATGCGACGAGCCCTCCCTCGAGACAAATT 175
? 56 efaTgLeuProAlaValGInclnUTrpGlyAlaGInclnProValHisLeu 72
? 176 CC.....ACACATGCTTTCACAAATGA.....AGCCTC 204
? 73 AspSerProAlaIleIyHisGInPheLeu.....LeuThGcIyAspTh 87
? 205 ATCTCAAGCAGCGCTCGAGCTGACTTCACTTATGACGTGCGACAGTCGACA 254
? 87 tGInGlyArGTyTArGcYAsGserGcIyLeuSerThr..... 99
? 255 CAGTGGAGAGTACAGGTGCGACGACAAACCTCCCGACCCCTCAGTAGACCCG 304
? 100 .....GlyTTP***GInLeu 104
? 305 TCCAGCTAGAGTCAATATCGCTGCTGTGCTC 335

seq_name: /cgn2_6/plodata/2/lna/5B_COMB.seq:US-08-667-939A-17

seq_documentation_block:
? Sequence 17, Application US/08667939A
? Patent No. 5998166
? General Information:
? APPLICANT: LUD, Shun
? TITLE OF INVENTION: CD16-II VARIANTS
? NUMBER OF SEQUENCES: 25
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BROWDY AND NEIMARK
? STREET: 419 Seventh Street, N.W., Suite 300
? STATE: Washington
? COUNTRY: USA
? ZIP: 20004
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? FILING DATE: 24-JUN-1996
? APPLICATION NUMBER: US/08/667-939A
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/433,123
? FILING DATE: 03-MAY-1995
```

```
?
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWDY, Roger L.
? REGISTRATION NUMBER: 25,618
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-628-5197
? TELEFAX: 202-737-3528
? INFORMATION FOR SEQ ID NO: 17:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 648 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE: CDS
? LOCATION: 1..645
? US-08-667-939A-17

alignment_scores:
? Quality: 70.50 Length: 79
? Ratio: 1.808 Gaps: 6
? Percent Similarity: 49.367 Percent Identity: 34.177

alignment_block:
? US-09-471-276-831 x US-08-667-939A-17 ..
?
? Align seg 1/1 to: US-08-667-939A-17 from: 1 to: 648
?
? 43 GIVGIVCysAspAlaAspValProclY.....ProProclY Asps 56
? 69 GGNACATGTGACTCTGAATGCGACGAGCCCTCCCTCGAGACAAATT 118
? 56 efaTgLeuProAlaValGInclnUTrpGlyAlaGInclnProValHisLeu 72
? 119 CC.....ACACATGCTTTCACAAATGA.....AGCCTC 147
? 73 AspSerProAlaIleIyHisGInPheLeu.....LeuThGcIyAspTh 87
? 148 ATCTCAAGCAGCGCTCGAGCTGACTTCACTTATGACGTGCGACAGTCGACA 197
? 87 tGInGlyArGTyTArGcYAsGserGcIyLeuSerThr..... 99
? 198 CAGTGGAGAGTACAGGTGCGACGACAAACCTCCCGACCCCTCAGTAGACCCG 247
? 100 .....GlyTTP***GInLeu 104
? 248 TCCAGCTAGAGTCAATATCGCTGCTGTGCTC 282

seq_name: /cgn2_6/plodata/2/lna/5B_COMB.seq:US-08-667-939A-12

seq_documentation_block:
? Sequence 12, Application US/08667939A
? Patent No. 5998166
? General Information:
? APPLICANT: LUD, Shun
? TITLE OF INVENTION: CD16-II VARIANTS
? NUMBER OF SEQUENCES: 25
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BROWDY AND NEIMARK
? STREET: 419 Seventh Street, N.W., Suite 300
? STATE: Washington
? COUNTRY: USA
? ZIP: 20004
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/667-939A
```

seq.documentation.block:
Sequence 16, Application US/0865793A
Patent No. 599816
GENERAL INFORMATION:
APPLICANT: LEO, Shun
TITLE OF INVENTION: CD16-II VARIANTS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
STREET: BROADY AND NEIMARK
ADDRESS: 419 Seventh Street, N.W., Suite 300
SPRINT, Washington
COUNTRY: D. USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release 11.0, Version #1.30


```

;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
seq_documentation_block:
; Sequence 2, Application US/09345468
; Patent No. 6245527
; GENERAL INFORMATION:
```



```

58 LeuprolaValIcInLUTrPGIyAlaIngInLuproVal.HisLeuAsp 74
167 GAGCCCGGCTGCGAGCGACCTCGGCGCTGAGACCTGACCGCTGGAGA 216
74 cr.....ProIa 76
217 AGCTGACTCCAGCAGCGTACAGCAGCAGCTGCTTCCTCCGAGC 266
77 11eLysIscIInPheLeuLeuLeuHrthcLysPheIcInLysIcIyAArg 93
267 ATGAGAGAGAGCTG.....GCTGAGAGCTGACCGCTG 298
93 sArgSerCylLeuSerThrLcIyTrp**G1eLueSerCylLeuLeuL 110
299 CTCCTACACAGACGAGACCTGCTGCTGCTGCGCCAGCAGCAGCTGAGC 348
110 eurthcLysProLysValLeuAla 117
349 TCGTTCACGCGAGAGTTTTCCTC 371
seq_name: /cgn2_6/p1data/2/lna/5h_COMB.seq:us-07-880-913-1
seq_documentation_block:
/ Sequence 1, Application US/07880913
/ Patent No. 5292651
/ GENERAL INFORMATION:
/ APPLICANT INFORMATION: ELLEN P.
/ APPLICANT: VAN COTT, ELIZABETH M.
/ APPLICANT: TARNON, CHRISTOPHER H.
/ TITLE OF INVENTION: METHOD FOR PRODUCING THE MAEI
/ TITLE OF INVENTION: RESTRICTION ENDONUCLEASE AND METHYLASE
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: 1427 S. 10TH ST.
/ STREET: 130 WATER STREET
/ CITY: BOSTON
/ STATE: MASSACHUSETTS
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ APPLICANT: ELLEN P.
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/880_913
/ FILING DATE: 04/13/93
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WILLIAMS, GREGORY D.
/ REGISTRATION NUMBER: 30901
/ REFERENCE/DOCKET NUMBER: 41606
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 523-9400
/ TELEFAX: (617) 523-9440
/ TUNING: 100.251 MHz
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3664 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ FEATURE:
/ NAME: CDS
/ LOCATION: 146..1099
/ US-07-880-913-1
alignment_scores:
Quality: 69.00 Length: 106
Ratio: 1.150 Gaps: 7

```

```

Percent Similarity: 56.604 Percent Identity: 33.019
alignment_block:
US-09-471-276-831 x US-07-880-913-1 ..
Align seg 1/1 to: US-07-880-913-1 from: 1 to: 3664
15 TrpLysProValTrpThrGluAlaAlaIlePheArgCylLysTrpIle**SerIe 31
2191 TGGGGTCTCTCACTCGGCGCGG.....TTTCGATGCTGCTGCTCT... 2232
31 utrPAlaIsSerCylHis**LeuLysThrIcInCylIncAsp.... 46
2233 ....GCAACAGAGCTGCGCCGCTATTTCAGCTGCTGCGAGCCGACCGT 2278
47 .....AlaAspValLProGlyProGlyLysPheSerArg 57
2279 GCGGACCCAGCGTGGAGAGAGCTTGAGACGCTGATGCTGCGCGAGCAG 2328
58 Leu ProAlaValIcInLUTrPGIyAlaIngInLuproVal.HisLeuAsp 74
2329 CTGGAGAGGTGCGGAGAGTGGGCGGCTGAG.....GCAACAGAGCTG 2372
74 erProAlaIle.....LysHisIscIInPheLeuLeuHrthcLys 85
2373 GACCAACAGCTGCTGCTGCTGCTCCAGAGAAACAGCTGAGAGCTGACCTGCGC 2422
86 AspThrcIcInLysArgTrp ArgCylLeuSerCylLeuSerThrCylLysTrp* 102
2423 CCAGCTGCGCGAGAGGCGGTGGCGAGAGCTGCGGTGCTGACGAGCAGTGGG 2472
102 **G1eLueSerCylLys 106
2473 AGCTGCTGAGCGCGC 2486
seq_name: /cgn2_6/p1data/2/lna/5h_COMB.seq:us-08-928-615-1
seq_documentation_block:
/ Sequence 1, Application US/08928615
/ Patent No. 5965535
/ GENERAL INFORMATION:
/ APPLICANT: Chaix, Pascal
/ APPLICANT: Stroobant, Vincent
/ APPLICANT: van der Bruggen, Pierre
/ TITLE OF INVENTION: NAC3-3 PEPTIDES PRESENTED
/ TITLE OF INVENTION: BY HLA CLASS II MOLECULES
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Wolf, Greenfield & Sachs, P.C.
/ STREET: 600 Atlantic Avenue
/ CITY: Boston
/ STATE: MA
/ COUNTRY: U.S.A.
/ ZIP: 02210-2211
/ COMPUTER READABLE FORM:
/ APPLICANT: Chaix, Pascal
/ OPERATING SYSTEM: IBM PC COMPATIBLE
/ SOFTWARE: PASTISO for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/928_615
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Chaix, Pascal
/ REGISTRATION NUMBER: 40122
/ REFERENCE/DOCKET NUMBER: 10463/7017
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-720-3500
/ TELEFAX: 617-720-2441
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:

```

```

:   LENGTH: 4204 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA
:   ORGANISM: HUMAN
:   ANTI-SENSE: NO
:   FEATURE:
:   NAME/KEY: Coding Sequence
:   LOCATION: 2465...3406
:   US-08-928-615-1

```

```

alignment_scores:
  Quality: 69.00      Length: 79
  Ratio: 1.917      Gaps: 3
  Percent Similarity: 45.570  Percent Identity: 29.114

```

alignment_block:
US-09-471-276-831 x US-08-928-615-1 ..

Align seg 1/1 to: US-08-928-615-1 from: 1 to: 4204

```

46 AspHisAspValProGlyProProGlyAspSerArgLeuProAlaValG1 62
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1100 GATGCGAGAC...CCACTCAGCAGGGGGCTGGGGCCGAGCCCTGCGAGGA 1146
62 nGluTrpGlyAlaGlnGluProValHisLeuAspSerProAlaIleLysR 79
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1147 GCGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165
79 LeuGlnPheLeuLeuThrGlyAspThrGlnGlyArgTyrFArgCysArgSer 95
|||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1166 .....GGAGGAGCTGAGCGGAGCTTGAGTCCAGATCA 1198
96 GlyLeuSerThrGlyTyr**GlnLeuSerLysLeuLeuGlnLeuThrG1 112
|||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1199 GTGGCAACTTGCGCTGGGGGATGCTGGGC.....ACAGT 1233
112 yProLysValLeuAlaCysSerLeuAlaLeuAspGly 124
|||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1234 GGCCAAATGTGCTGTGTGCTCATGTGCGCTTCAGAGCT 1270

```

•
•
•

KM gene therapy: chromosome mapping; upstream regulatory sequence;
 KM forensic: location; development; protein synthesis; stability;
 KM regulation: identification; ss.
 XX Homo sapiens.
 CN W09953051-A2.
 XX
 PD 21-OCT-1999.
 XX
 XX 09-APR-1999: 99MO-1800712.
 XX
 PR 09-APR-1998: 98US-0057719.
 PR 28-APR-1998: 98US-0069047.
 XX
 XX (GEST) GENSET.
 PA
 XX Dunas Mline Edwards J, Duclert A, Giordano J;
 XX WPI: 2000-038446/03.
 DR P-PSDB: AAY64670.
 XX
 PT Novel secreted protein 5', expressed sequence tag sequences used in
 XX diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX
 XX Claim 1: Page 195: 837bp: English.
 XX
 CC AA42265 to AA423075 represent novel 5', expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY65438 represent the EST-related proteins corresponding to AA42265 to
 CC AA423075. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC expression, development, stage, rate, and quantity of protein synthesis, as
 CC well as the development of a disease. The ESTs can be used to identify
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the
 CC use of a polypeptide as a membrane or membrane-anchored polypeptide
 CC into a cell. The peptides encoded by the ESTs are useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AA42265 to AA42264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 XX
 XX Sequence 412 BP: 79 A: 131 C: 119 G: 80 T: 3 other:
 SO
 Alignment_scores:
 Quality: 655.00 Length: 126
 Ratio: 5.325 Gaps: 0
 Percent Similarity: 97.619 Percent Identity: 97.619
 Alignment_block:
 US-09-471-276-831 x AA42264 ..
 Align seg 1/1 to: AA42284 from: 1 to: 412
 1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrPdp1p: 17
 |||||||
 35 ATGTCACATGCTGCGCTTCTCTGTCGGCTGCTACACCTGGGCC 84
 |||||||
 17 GVALThGtLualalalaj1lePhyTrGtLthGtL***SeclatTfGalic 34
 |||||||
 85 AGTACACGAAGCACACCATATTTTATGAGCCGACGACCCATGCGGAC 134
 |||||||
 34 luseTcLlulllull***leuTyThrLeuclYgInGyAsAPAlaAspVlpro 50
 |||||||
 135 AGTCCGAAACACTGAGTGAAGAACCTCTGGCCCAATGTCACCTGACGTGCA 184

51 G1pTProdiLyaSepcArGyleuProAlaValGInG1uTfPpG1YAlaG1 67
 |||||||
 185 GCGCCGCCCTGGAGACTCCAGACTCCAGCTCCATCAAGAAATGGGGGCCCA 234
 |||||||
 67 nGtUroValAlHlStenAAspSPROAlal1clYgInG1InPheLeuclat 84
 |||||||
 235 GAGCCCTGTGCACCTTCACTACCTGCCATCCAGCCAGCCAGCTTCTGTGA 284
 |||||||
 84 hrc1YAsPThrGInG1uTfYtArGyArGseArGyluAsSerThrcY 100
 |||||||
 285 CGGTGACACCCAGGCGCCCTACCCGCGCCCTCGGCTTGTCCACAGCA 334
 |||||||
 101 TrP***GtLleSerTleuLeuclYlulGThGtYpAdLysyLleuAl 117
 |||||||
 335 TCGKCCAGCTGACCAACCTCTGAGACCTGACAGCGCCAAAAGTCTTCC 384
 |||||||
 117 acYSerTleuAlaIeuaSPG1YAlaSer 126
 |||||||
 385 CTCCTCTCCCTGCTTCGATGCGCCCACT 412
 seq_name: /STDS2/gcgcdata/geneseq/geneseq/MA1999.DAT:AAV88297
 seq_documentation_block:
 ID AAV88297 standard: cDNA: 276 BP.
 XX
 AC AAV88297:
 XX
 XX 12-FEB-1999 (first entry)
 DE EST clone GD177.
 XX
 KM Expressed sequence tag: secreted protein: haematopoiesis regulator;
 KM tissue growth; activating; inhibin; tumour invasion suppressor; EST: human;
 KM chemokinesis; chemokinesis; haemostasis; gene therapy; Chromolysins;
 KM receptor; ligand: anti-inflammatory; tumour inhibitor; ds.
 XX
 XX Homo sapiens.
 XX
 XX W09845437-A2.
 XX
 PD 15-OCT-1998.
 XX
 PD 10-APR-1998: 98MO-US06956.
 XX
 PR 10-APR-1997: 97US-0837312.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D;
 XX Bacle LA, Spaulding V, Treacy M;
 XX WPI: 1999-070078/06.
 DR
 XX
 XX New polynucleotides encoding human secreted proteins - derived from
 XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 XX ovary, pituitary, retina and colon cDNA libraries
 PT
 XX
 XX Claim 1: Page 348-349: 641pp: English.
 XX
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences and proteins isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include: haematopoiesis regulating activity, tissue growth activity,
 CC haemostasis regulating activity, chemokinesis/chemokinetic activity,
 CC activating/inhibin activity, receptor/ligand activity, anti-inflammatory
 CC activity, chemotactic activity, receptor/ligand activity, anti-inflammatory
 CC activity, chemotactic activity, receptor/ligand activity, anti-inflammatory
 CC activity, chemotactic activity, receptor/ligand activity, anti-inflammatory
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.

XX PS Claim 1: Page 51-52; 63pp; English.

XX The invention relates to a method of diagnosing abnormal levels of
XX growth hormone (GH) activity in the liver or predicting a change in the
XX condition of the liver in response to abnormal GH activity. The method
XX involves correlating the level of expression of certain specific genes
XX with the level of GH activity in the liver, or with an expected change
XX in the condition of the liver as the result of GH activity. Excessive GH
XX activity in the liver is thought to be deleterious to health, causing an
XX increase in liver size as a consequence of both hyperplasia and
XX hepatocyte hypertrophy and hepatocellular carcinoma. The process with
XX which the liver size increases is called fatty liver disease, or non-
XX growth hormone identified a number of genes whose expression in the
XX liver is altered by high GH levels. The genes which are upregulated
XX those encoding alpha-fetoprotein, corticosteroid binding globulin,
XX fetuin, rat-8-interacting protein, paroxanase-3, cytochrome P45011A,
XX S-2 hydroxyacid oxidase, interferon alpha/beta receptor, growth hormone
XX receptor, prolactone 2-subunit, and coagulation factor V. Two novel
XX genes, clone 5 (MAG8725) and clone 43 (MAG8726) are also upregulated
XX in the liver in response to abnormal levels of GH activity. The level of the
XX gene encoding 3-beta-hydroxysteroid dehydrogenase/delta-5-delta-4
XX isomerase appears to be downregulated. The invention also relates to
XX transgenic nonhuman mammals comprising a GH-inducible transgene which
XX exhibit or have a propensity to develop a liver pathology, an assay for
XX drugs which inhibit the development of, or which treat a liver pathology
XX comprising administering the drug to the transgenic animal; and
XX preventing or treating a liver pathology in a patient consisting
XX of administering a GH-inhibiting agent to a patient with GH and
XX disease. The method of the invention is used for diagnosing abnormal levels of GH
XX activity in the liver or predicting a change in the condition of the
XX liver in response to abnormal levels of GH activity. The GH-regulatable
XX liver genes and proteins are useful as diagnostic markers of liver
XX pathology. Assays for the expression of these genes is useful for the
XX diagnosis of liver pathologies associated with gylactem or acromegaly.
XX The invention also relates to a transgenic mouse model of liver disease
XX upon the same genes. The present sequence represents the novel mouse
XX liver cDNA clone 5, expression of which is upregulated by abnormal GH
XX levels.

XX Sequence 1908 BP: 525 A; 475 C; 477 G; 431 T; 0 other:

XX alignment_scores:

Quality:	139.50	Length:	138
Ratio:	1.744	Gaps:	4
Percent Similarity:	57.971	Percent Identity:	29.710

XX alignment_block:

XX 05-09-471-276-831 x MAG87257 ..

XX Align seg 1/1 to: MAG87257 from: 1 to: 1908

```
1 MelsrMGLnauValValPheUleuLeuTrpGlyValTrpPLeuLP 17
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
11 AATCTCTCGAGCTACAGTACGATCTGCGCTCGGAGGTCACTGCGGCC 60
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 ValThrGlnAlaAlaIlePheTrpGLeuThrGln***SerLeuTrpAla 34
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 AGGAAAGACTGATTCGATCTGGCGAGAGACAGTCAAGTCAATGAGAG 110
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
111 AGCTCAGATCGCTGGGAAACCTGGGCAAACTGACCTGGGTGTTCAC 160
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
47 AlaaPValPLeuProGlyPLeuSerArgLeuTrpAlaValaInoG1 63
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 GTTATGTTTCCGCG.....ACTAGATCTTCAGCTACAGTCAAGAG 198
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 uTrpPValaG1nG1uPValaValaIleuLeuPLeuSerProAlaIleuYn 80
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 CGAGTGGTTCCTGAGTCAAGTCTCAAGTCAAGACACAGTGTGTGTATCC 248
```

```

80 1nPhleuLeu.....ThrGlyAspThrGlnGlyArgTyrArgCys 92
   ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
249 CcTrrTCcCTGGGGCCATTACAGTACAGACAGTGCACCTGCCCTGC 298
   ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 ArgSerGlyLeuSer.....ThrG1 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 AGnTrrGCTGGACACCCCTGTCGATTCACCTGCCACGACGTGACAA 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 YTrp**G1nLeuSerLysLeuGlnGluThrGlyProLysValLeuA 117
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
349 CTrGACACATCTTCATGCATCTGTCGAGCTACAGGAAAGGCCCTTC 398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 LacSerLeuAla 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 CTCGCCCTCTTCCT 412

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT:AA502105
seq_documentation_block:
AA502105 standard: cDNA; 1477 BP.
AA502105:
18-JUL-2001 (first entry)
Human MANCO 511, variant #4 cDNA sequence.
Human: MANCO 511; transmembrane protein; diagnostic; asthma;
immunological disorder; arthritis; graft rejection; renal disorder;
acquired immunodeficiency syndrome; inflammatory disorders; psoriasis;
AIDS; embryonic disorder; brain; cerebral oedema; ischaemia; tumour;
prostate; cerebrovascular disease; pituitary; Cushing's disease;
neurodegenerative disease; Parkinson's disease; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX PH 108..1007
XX CDS /*tag" a
XX /product: "MANCO 511, variant #4"
XX
XX MO200121631-A2.
XX
XX 29-MAR-2001.
XX
XX 20-SEP-2000: 2000MO-US25982.
XX
XX 20-SEP-1999: 99US-0399723.
XX
XX (MILL-) MILENITUM PHARM INC.
XX
XX Kirtz SJ, Sharp JD, Fraser CC, Barnes T, Kingsbury G:
XX WPI: 2001-211461/21.
XX P-PSDB: AAUD1394.
XX
XX New nucleic acid encoding INTERCEPT 307, MANCO 511, TANCO 351, TANCO
XX 352, MANCO 308 secreted or transmembrane protein, useful
XX for the diagnosis and treatment of arthritis, psoriasis and Parkinson's
XX disease.
XX
XX Disclosure: Page 300-301; 362pp; English.
XX
XX The sequence represents the coding sequence of human MANCO 511
XX variant #4 transmembrane protein. The nucleic acid and polypeptide
XX sequences have been deposited with the National Center for Human
XX Immunological Disorders (e.g. arthritis, graft rejection, and acquired
XX immunodeficiency syndrome), inflammatory disorders (e.g. psoriasis and
XX asthma), renal disorders, embryonic disorders, brain-related disorders
XX (e.g. cerebral oedema), cerebrovascular diseases (e.g. ischaemia),
XX tumours, prostate-related disorders, pituitary-related disorders (e.g.
XX Cushing's disease) and neurodegenerative diseases (e.g. Parkinson's
XX disease).

```

```

XX
XX SO Sequence 1477 BP; 393 A; 390 C; 372 G; 322 T; 0 other:
XX
XX alignment_scores:
XX      Quality: 109.00      Length: 120
XX      Ratio: 1.627
XX      Percent Similarity: 55.833      Percent Identity: 34.167
XX
XX alignment_block:
XX US-09-471-276-831 x AA502105
XX
XX Align seg 1/1 to: AA502105 from: 1 to: 1477
XX
XX 1 MetSerMetLeuValVal1nPhleuLeuLeuTrpGlyValThrTrpGlyPr 17
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 162 GTGAGCCCTCCCTCCATGCTCTTCTGCTGCTGCGCGAGGTGGGCGC 211
XX
XX 17 oValThrGlnAlaAla1eIlePheTyrGluThrGln***SerLeuTrpAlaG 34
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 212 CAGAGCCAGGTGCAGCGACGAGACACCTGCGCAACGACGCTGGAGTCG 261
XX
XX 34 luseTlueHis**1ealuYThrLeuGlyGln..CysAspAlaAspValP 50
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 262 AGCCAGAGCTCTGTGATC...AGCGGGGAGACCTGTGATCACCTCGAGTCT 308
XX
XX 50 TrcGlyProGlyAspSerArgLeuProAla..ValGlnGlu..... 63
XX
XX 309 CAGGGACCTCGAGCGCCCGACGATACGCTCTGCTTAAAGAGGAGACCC 358
XX
XX 64 .....TrpGlyAlaGlnGluProVal1n1ealuAspSerProAla1leI 78
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 359 AGAACCCTGGGACACAGAGACGACCACTG.....GACCCCAAGACA 399
XX
XX 78 YnH1sG1nPhleuLeuLeuThrGlyAspThrGln.....GlyArgTyr 91
XX
XX 400 AGCCAGATTCCTCCATCCCATCCATCCAGACAGACGACCGTGCAGAGATAC 449
XX
XX 92 ArgCysArgSerGlyLeuSerThrGlyTrp**G1nLeuSerLysLeuG 108
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 450 GCCTGTTACTACTACAGCCCTCGACGCTGTCTCAGAGCTCAGGCCCTC 499
XX
XX 108 nCULeu 110
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 500 GGAAGCTG 506
XX
XX seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AA54610
seq_documentation_block:
ID AA54610 standard: cDNA; 956 BP.
AA54610:
11-APR-2001 (first entry)
Leukocyte Immunoglobulin like receptor coding sequence (LIR-9s1).
Leukocyte Immunoglobulin like receptor, LIR; gene therapy;
autoimmunity; autoimmune disorders; immune system; human; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX PH 115..912
XX CDS /*tag" a
XX /product: "Leukocyte Immunoglobulin like receptor
XX
XX MO200068383-A2.
XX
XX 16-NOV-2000.
XX
XX 12-MAY-2000: 2000MO-US13228.
XX
XX

```

```

PR 12-MAY-1999; 9905-0310463.
XX (IMNV ) IMMUNEX CORP.
XX
XX Cosman DJ, Anderson DM, Borges L;
XX MPI: 2000-587645/67.
XX P-PSDB: AAB04180.
XX
XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
XX treating autoimmune diseases and disease states with suppressed immune
XX function
XX
XX Claim 2: page 109-110; 117pp; English.
XX
XX Host cells transformed with a vector which are capable of
XX expressing a leukocyte immunoglobulin like receptor (LIR), can be
XX used to produce the LIR polypeptide. LIR coding sequences may be
XX used in the gene therapy of disorders mediated directly or
XX indirectly by defective or insufficient amounts of any of the LIR
XX polypeptides. The LIR polypeptides can be used to treat autoimmune
XX diseases and disease states with suppressed immune function.
XX
XX Sequence 956 BP; 220 A; 291 C; 269 G; 176 T; 0 other:

alignment_scores:
Quality: 108.00 Length: 120
Ratio: 1.612 Gaps: 6
Percent Similarity: 55.833 Percent Identity: 34.167

alignment_block:
US-09-471-276-831 x AAAS4610 ..

Align seq 1/1 to: AAAS4610 from: 1 to: 956

1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGlyPyr 17
169 GTGAGCCCTGCCCTCATGGTTCTGCTCTGCTGCGCTGACGTGGGCC 218
17 oValThrGluAlaAla1LePheTyrGluThrGln**SerLeuTrpAlaG 34
219 CAGGACCCACGTCGACGACGAGCAACCTCTCCAAAGCCACCTCTGGGCG 268
34 lUseSerLuhHis**LeuLysThrLeuGlyGln..CysAspAlaAspAlp 50
269 AGCCGACGCTCTGTATC...AGCCGGGAGACACGTGTGACATCCGCTCT 315
50 roGlyProProGlyAspSerArgLeuProAla..ValGlnGln..... 63
316 CAGGAGACCTGCGACCCCAAGACATACCTCTGCTGCGAGGAGAGAGCC 365
64 .....TrGlyAlaGlnGlnGlyProValHisLeuAspSerProAlaLe 78
366 AGAACCCCTGGGACACAGAACCCCTG.....GACCCCAAGACACA 406
78 yHisLscLpPheLeuLeuThrGlyAspThrGln.....GlyArgTyr 91
407 AGCCAGACATCTCCATCCCATCCACACAGACACCCATGACGACATAC 456
92 ATGcYtAAgSerGlyLeuSerThrGlyTrp**GlnLeuSerLysLeuLe 108
457 CGCTGTACTACTAGACGCTCGAGCTGTGTGACAGCCGACGACCCCT 506
108 uGluLeu 110
507 GGAAGCTG 513

seq_name: /STD52/cgdata/geneseq/geneseq/MA2000.DAT.AAAS4608
seq_documentation_block:
ID AAAS4608 standard; cDNA: 1016 BP.
XX

```

```

AC AAAS4608:
XX
XX 11-APR-2001 (first entry)
XX
XX Leukocyte immunoglobulin like receptor coding sequence (LIR-9m1).
XX
XX Leukocyte immunoglobulin like receptor; LIR; gene therapy;
XX autoimmunity; autoimmune disorders; immune system; human; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 69..968
XX /tag="a
XX /product=" Leukocyte immunoglobulin like receptor
XX
XX WO200068383-A2.
XX
XX PD 16-NOV-2000.
XX
XX 12-MAY-2000; 2000MO-US13228.
XX
XX 12-MAY-1999; 9905-0310463.
XX
XX (IMNV ) IMMUNEX CORP.
XX
XX Cosman DJ, Anderson DM, Borges L;
XX MPI: 2000-587645/67.
XX P-PSDB: AAB04180.
XX
XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
XX treating autoimmune diseases and disease states with suppressed immune
XX function
XX
XX Claim 2: Page 104-105; 117pp; English.
XX
XX Host cells transformed with a vector which are capable of
XX expressing a leukocyte immunoglobulin like receptor (LIR), can be
XX used to produce the LIR polypeptide. LIR coding sequences may be
XX used in the gene therapy of disorders mediated directly or
XX indirectly by defective or insufficient amounts of any of the LIR
XX polypeptides. The LIR polypeptides can be used to treat autoimmune
XX diseases and disease states with suppressed immune function.
XX
XX Sequence 1016 BP; 231 A; 310 C; 282 G; 193 T; 0 other:

alignment_scores:
Quality: 108.00 Length: 120
Ratio: 1.612 Gaps: 6
Percent Similarity: 55.833 Percent Identity: 34.167

alignment_block:
US-09-471-276-831 x AAAS4608 ..

Align seq 1/1 to: AAAS4608 from: 1 to: 1016

1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGlyPyr 17
169 GTGAGCCCTGCCCTCATGGTTCTGCTCTGCTGCGAGCTGAGTGGGCC 172
17 oValThrGluAlaAla1LePheTyrGluThrGln**SerLeuTrpAlaG 34
219 CAGGACCCACGTCGACGACGAGCAACCTCTCCAAAGCCACCTCTGGGCG 222
34 lUseSerLuhHis**LeuLysThrLeuGlyGln..CysAspAlaAspAlp 50
223 AGCCGACGCTCTGTATC...AGCCGGGAGACACGTGTGACATCCGCTCT 269
50 roGlyProProGlyAspSerArgLeuProAla..ValGlnGln..... 63
270 CAGGAGACCTGCGACCCCAAGACATACCTCTGCTGCGAGGAGAGCC 319

```

[illegible]

[illegible]

```

seq_name: /SIDSG/cgdata/geneseq/geneseqn/NM2001.DMT:AAS02103
seq_documentation_block:
ID ID AAS02103 standard: cDNA: 1477 BP.
CC AAS02103:
XT XT 18-Jun-2001 (first entry)
DE DE Human MANGO 511, variant #2 cDNA sequence.
XX XX
XX Human: MANGO 511; transmembrane protein; diagnostic; asthma;
KW immunological disorder; arthritis; graft rejection; renal disorder;
KW acquired immunodeficiency syndrome; inflammatory disorders; psoriasis;
KW neurofibromatosis; disease; cerebral oedema; ischaemia; tumour;
KW neurodegenerative diseases; Parkinson's disease; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 108..1007
FT "tag" =
FT /product= "MANGO 511, variant #2"
PN NM020012.631-AZ.
XX
XX 29-MAR-2001.
PD PD
PF 20-SEP-2000; 2000MO-U525582.
XX
XX 20-SEP-1999; 990US-0359723.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI Kirst SJ, Sharp JD, Fraser CC, Barnes T, Kingsbury G;
DR WPI: 2001-211461/21.
P-PDB: AA001392.
XX
XX New nucleic acid encoding INTERCEPT 307, MANGO 511, TMANGO 351, TMANGO
XX 307, TMANGO 351, TMANGO 307, TMANGO 351, TMANGO 307, TMANGO 351, TMANGO
XX for the diagnosis and treatment of arthritis, psoriasis and Parkinson's
XX disease -
XX
XX Disclosure: Page 296-297; 362pp: English.
XX
XX The sequence represents the coding sequence of human MANGO 511
XX variant #2 transmembrane protein. The nucleic acid and polypeptide
XX sequences are useful for the diagnosis, prognosis and treatment of
XX immunodeficiency syndrome, inflammatory disorders (e.g. psoriasis and
XX asthma), renal disorders, embryonic disorders, brain-related disorders
XX (e.g. cerebral oedema), cerebrovascular diseases (e.g. ischaemia),
XX tumours, prostate-related disorders, pituitary-related disorders (e.g.
XX Cushing's disease) and neurodegenerative diseases (e.g. Parkinson's
XX disease).
XX
XX Sequence 1477 BP; 393 A; 390 C; 372 G; 322 T; 0 other:
XX

```

see documentation block:

ID AAV44827 standard: DNA: 1550 BP.
 AC AAV44827:
 XX 20-OCT-1998 (first entry)
 DE FcR-IV coding sequence.
 XX FcR-IV coding sequence.
 XX Fc receptor-like protein; phagocytosis inducer; rheumatoid arthritis;
 XX immune complex related disease; systemic lupus erythematosus; allergy;
 XX haemolytic anemia; thrombocytopenia; anaphylaxis; cancer; lymphoma;
 XX Leukemia; Infection; Immunomodulator; viral entry inhibitor; therapy;
 XX FcR-IV ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FH CDS 22..1440
 FT /tag- a
 FT /product= FcR-IV
 FT sig_peptide 22..1439
 FT met_peptide 70..1437
 FT /tag- c
 XX MO9831806-A2.
 XX 23-JUL-1998.
 XX 20-JAN-1998: 98MO-US01184.
 XX 18-JUN-1997: 9705-0049872.
 XX 21-JAN-1997: 9705-0034205.
 XX (BDNA-) HUMAN GENOME SCI INC.
 XX Gentz RL, Murphy M, N1 J, Olsen HS, Ruben SM;
 XX WPI: 1998-414105/35.
 XX P-PSDB: AAM69234.
 XX Nucleic acid encoding Fc receptor-like polypeptides or their
 XX fragments - and related vectors, transformed cells and antibodies,
 XX useful for creating and diagnosing diseases of the hematopoietic
 XX and immune systems
 XX Claim 2: Fig 4A; 141pp: English.
 XX This sequence encodes the Fc receptor-like IV protein (FcR-IV) of the
 XX invention. Cells containing the DNA are used to express the recombinant
 XX protein, and to screen for specific (ant)agonists. The proteins are used
 XX to induce phagocytosis, and chelate (ant)agonists are used to treat immune
 XX complex related diseases, systemic lupus erythematosus, allergy, anaphylaxis,
 XX haemolytic anemia, thrombocytopenia, lymphoma, leukemia, infection;
 XX allergy, colorectal or breast cancer, lymphoma, leukaemia, infection by
 XX intercellular pathogens etc.). The antagonists are also useful as
 XX immunomodulators and inhibitors of viral (e.g. human immune deficiency or
 XX dengue virus) entry into cells. The proteins may also be used to screen
 XX for specific binding agents, i.e. (ant)agonists, for raising antibodies
 XX (ant)agonists, and chelate (ant)agonists are used to treat immune
 XX complex related diseases, systemic lupus erythematosus, allergy, anaphylaxis,
 XX haemolytic anemia, thrombocytopenia, lymphoma, leukemia, infection;
 XX demyelination of the levels of expression of the proteins and for
 XX affinity purification of the proteins. The DNA and its fragments are
 XX useful as hybridisation probes or primers for isolating genes, in
 XX situ hybridisation (chromosome mapping) and diagnostically to measure
 XX mRNA expression.
 XX Sequence 1550 BP; 363 A; 506 C; 414 G; 267 T; 0 other;

alignment_scores:
 Quality: 104.50 Length: 121
 Ratio: 1.633 Gaps: 7

Percent Similarity: 52.893 Percent Identity: 36.364
 alignment_block:
 US-09-471-276-831 x AAV44827 ..
 Align seg 1/1 to: AAV44827 from: 1 to: 1550
 2 SerMetLeuValValPhe.....LeuLeuLeuThrPro1ValThrTrpG1 16

 19 GCCATGATCCCACTTCAGCGCTCTGCTCCGCGCCGCGATGTGG 68
 16 pProValThrGlnAlaAlaIlePheTyrGlnThrGln***SerLeuTrpA 33

 69 CCGCAGTCCCACTTCAGCGCTCTGCTCCGCGCCGCGATGTGG 118
 33 IacLysSerCysHis**LeuValThrLeuGlyGln.....CysAspAlaAsp 48

 119 CTGACGAGGCTCTGTGATACG....CTGGGGGACCTCTGTACCATCTCT 164
 49 ValProGlyProProDilyAspSerArgLeuProAla ValGlnGln.... 63

 155 GTTCAGAGACCTCTGAGCTCTGAGAGTACCTCTGAGATAGAGAGAA 214
 64TrpGlyAlaGlnGlyProValHisLeuAspSerProAla 76

 215 GCCCAGCACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255
 77 IleLysHisIleAspPhe.....LeuLeuThrGlyAspThrGlnGlyArg 90

 256 AACAGAGCGAGTCTCTGATCCATCCATCCAGAGAGAGAGAGAGAGAG 305
 50 GTTArgCysArgSerGlyLeuSerThrGlyTyrPr**GlnLeuSerTyrSL 107

 306 ATACCCCTGTACTATCCAGGACCTCTGTAGCTGTATCCAGGACCATGTGAC 355
 107 euLeuGlnLeu 110
 356 CCTTGAAGACTG 366
 seq_name: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT:AAV69337
 seq_documentation_block:
 ID AAV69337 standard: DNA: 1625 BP.
 AC AAV69337:
 XX 09-FEB-1999 (first entry)
 XX Human IIR-pbm2 cDNA.
 DE Human IIR-pbm2 cDNA.
 XX LIR-pbm2: immunoregulator; leukocyte immunoglobulin-like receptor;
 XX therapeutic; treatment; disorder; FcR; MHC class I; receptor; inhibitor;
 XX negative signaling; autoimmune disease; suppressor; IIR; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FH CDS 30..1376
 FT /tag- a
 FT /product= "LIR-pbm2"
 XX MO9848017-A1.
 XX 29-OCT-1998.
 XX 23-APR-1998: 98MO-US00244.
 XX 24-APR-1997: 9705-0842248.
 XX (IMM) IMMUNEX CORP..
 XX Cosman DJ;

XX WPI: 1998-609950/51.
 DR P-PSDB: AAM82551.

XX Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful
 for treating autoimmune diseases or disease states associated
 with suppressed immune function

PS Example 9; Page 83-85; 112pp: English.

XX This sequence encodes a novel leukocyte immunoglobulin-like receptor
 LIR polypeptide. LIR-pm2. This sequence can be administered
 to a patient to treat a disease state associated with suppressed
 immune function. LIR-pm2 is a member of the LIR family of proteins
 CC contains cytoplasmic immunoreceptor tyrosine-based inhibitory motifs
 CC (ITIMs). While other LIR family members lack ITIMs, by analogy with the
 CC structure and function of known MHC Class I receptor molecules, LIRs
 CC having ITIMs are inhibitory receptors mediating negative signaling.
 CC While those lacking ITIMs are activatory receptors. Failure of a
 CC LIR to express negative signaling could result in autoimmune
 CC diseases while failure of a receptor to express negative signaling
 CC could result in suppressed immune function. They are also useful to
 CC produce probes for detecting LIR nucleic acids or isolating LIR DNA from
 CC other species.

XX Sequence 1625 BP; 399 A; 523 C; 425 G; 278 T; 0 other;

Alignment_scores:
 Quality: 103.50 Length: 121
 Ratio: 1.617 Gaps: 7
 Percent Similarity: 52.893 Percent Identity: 36.364

alignment_block:

US-09-471-276-831 x AAV69337 ..

Align seg 1/1 to: AAV69337 from: 1 to: 1625

```

2 SerMetLeuValValPhe.....LeuLeuLeuTrpCylValThrTrpG1 16
:::||||| ::|||:| |::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
27 GCCATGATCCCGACCTTCACGCGCTGCTGCGCTGCGCGCTGATCTGGG 76
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
16 YPcValaThnGluAlaAlaIlePheTrpGln***SerLeuTrpA 33
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
77 GCCGAGACCATCATCGACGAGGCCCTGCCAACCACCTCTGGG 126
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
33 IaG1userGluIle***LeuLystrIleuGlyIle...CysAspAlaAsp 48
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
127 CTGACGACGCGCTCTGATCAG...CTGGGCAAGCTCTGTGATACCATCTG 172
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
49 ValProGlyProArgGlyAspSerTrpGluProAla..ValG1Glu... 63
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
173 GTTCAGAGGACCTCTGAGACCTGGGAGTACCTCTGATACCAAGAGAA 222
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
64 .....TrpCylValGlnGluProValIleLeuAspSerProAla 76
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
223 GCCCAGACACCTCTGGGACAGCAAGAACCCACATG.....GACCCCAAG 263
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
77 IleLystrIleGlnPhe.....LeuLeuTrpCylValThrTrpGlnIle 90
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
264 AACAGCCGACATTCATCCATCCATCATCGACAGAGGATCATACAGGAG 313
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
90 gTfYArGcYAsrGerglyLeuSerTrpG1Trp**GlnLeuSerLysL 107
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
314 ATACGCGCTGTACTACATCCAGCCCTGAGGCTGTGATACAGCCACGAG 363
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
107 euLeuGluLeu 110
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
364 CCCGAGACCTG 374
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

seq_name: /SI/SDA/geneseq/geneseq/NA1998.DAT:AAV23273

seq_documentation_block:

ID AAV23273 standard; cDNA; 1625 BP.
 XX AAV23273:

XX 17-JUL-1998 (first entry)

XX Human gp49 HM18 encoding cDNA.

XX Human: gp49; HM18; HM43: Immunoglobulin: Immune response: mast cell;

XX bone marrow; cell-surface member; FCER1; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS /tag.. a /product.. "gp49 HM18"

XX W08090638-A1.

XX 12-MAR-1998.

XX 05-SEP-1997; 97MO-US15586.

XX 06-SEP-1996; 96US-0025846.

XX (BGM) BRIGIAM & WOMENS HOSPITAL.

XX Alm JP, Austen KF, Castells MC, Katz HR:

XX WPI: 1998-193318/17.

XX P-PSDB: AAM53463.

XX Cell-surface member of immunoglobulin super-family, human gp49 -

XX useful to treat undesired immune responses, especially mast

XX cell-related diseases

XX Claim 5; Fig 1A-B; 62pp: English.

XX The present sequence encodes human gp49 polypeptide HM18. The present

XX Invention also describes: (1) a fusion polypeptide comprising a human

XX gp49 component and FCER1 or a detectable marker; (2) a recombinant

XX nucleic acid encoding human gp49 or a human gp49-FCER1 fusion as above;

XX and (4) an antibody which selectively binds to gp49. Membrane gp49 (1),

XX its related DNA can be used to treat an undesired immune response,

XX especially a mast cell-related disease.

XX Sequence 1625 BP; 400 A; 519 C; 429 G; 277 T; 0 other;

Alignment_scores:
 Quality: 103.50 Length: 121
 Ratio: 1.617 Gaps: 7
 Percent Similarity: 52.893 Percent Identity: 36.364

alignment_block:

US-09-471-276-831 x AAV23273 ..

Align seg 1/1 to: AAV23273 from: 1 to: 1625

```

2 SerMetLeuValValPhe.....LeuLeuLeuTrpCylValThrTrpG1 16
:::||||| ::|||:| |::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
43 GCCATGATCCCGACCTTCACGCGCTGCTGCGCTGCGCGCTGATCTGGG 92
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
16 YPcValaThnGluAlaAlaIlePheTrpGln***SerLeuTrpA 33
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
93 GCCGAGACCATCATCGACGAGGCCCTGCCAACCACCTCTGGG 142
:::|||||:|:|:~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
33 IaG1userGluIle***LeuLystrIleuGlyIle...CysAspAlaAsp 48
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
143 CTGACGACGCGCTCTGATCAG...CTGGGCAAGCTCTGTGATACCATCTG 188
:::|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```


alignment_scores:

Quality: 90.50 Length: 138
Ratio: 1.560 Gaps: 6
Percent Similarity: 42.029 Percent Identity: 28.986

alignment_block:

US-09-471-276-831 x AAV44826 ..

Align seg 1/1 to: AAV44826 from: 1 to: 1991

```

1 MetSerMetLeuValValPheLeuLeuThrProGlyValThrProGlyPro 17
73 ATGACGCCGCCCTCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
17 oValThGluAlaIlePheThyTrpGluThGln***SerLeuTrpAlaG 34
13 CAGGACCCGCGTGCAGAGGAGCCCTTCCCAACCCAGCCCTCGGGCTG 172
34 LuserGluHis***LeuLysThrLeuGlyGlnGlyncysAspAlaAspValPro 50
173 AG.....AlaGlnGlu..... 174
51 GlyProProGlyAspSerArgLeuProAlaValGlnGluTrpGly..... 65
175 .....CCAGGC.....TCTGTATACGCTGGCGGAGCC 203
66 .....AlaGlnGlu..... 68
204 CGTGACCATCTGTCGTGACGGAGGAGCCCTGAGAGCCGAGAGTACCACTG 253
69 .....ProValHisLeuAsp.....SerPro 75
254 ATAAAGAGGAGAGCCAGAGCCCTTGGAGACAAATACCCCACTGGAACCC 303
76 AlaIleLysHisGlnPheLeuLeuThrGlyAspThrGln.....G1 89
304 AAGAACAGAGGCCAGATTCCTCATCCATCCATGACAGACAGCCAGTCGAG 353
89 YALGTGYARGCYASRGSERGLYLSERSETHRGLYTRP**G1LUSERSERL 106
354 GAGATGACCGCTCCGACATTTACAGAGCTGTCAGAGCTGTCAGAGCCAGCG 403
106 YSLLeuLeuGlnLeu 110
404 ACCCCCTGGAGGCT 417

```

seq_name: /S1SD2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV69338

seq_documentation_block:

10 AAV69338 standard: DNA: 2194 BP.

AAV69338:

DT 09-FEB-1999 (first entry)

Human LIR-pbm17 cDNA.

LIR-pbm17: immunoregulator; leukocyte immunoglobulin-like receptor;
Therapeutic: treatment: disorder: ITTM; MHC class II receptor inhibitor;
cytoplasmic immunoreceptor tyrosine-based inhibitory motif; activator;
negative signalling; autoimmune disease; suppressor; LIR: ss.

OS Homo sapiens.

Key Location/Qualifiers
67 10633 /tag= a
CDS /product= "LIR-pbm17"

MO5848017-A1.
29-OCT-1998.

XX 23-APR-1998; 98WO-0508244.
XX 24-APR-1997; 97US-0642248.
XX (IMW) IMMUNEX CORP.
XX
XX Cosman DJ;
XX
XX WPI: 1998-60990/51.
XX P-PSDB: AAM82952.

XX Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful
XX e.g. for treating autoimmune diseases or disease states associated
XX with suppressed immune function
XX
XX Example 9; Page 87-89; 112pp: English.

XX This sequence encodes a novel leukocyte immunoglobulin-like receptor
XX (LIR) polypeptide LIR-pbm17. This sequence can be administered
XX orally or injected into a subject to elicit an immune response. A defective
XX amount of LIR polypeptide LIR-p32 is associated with the LIR family members
XX contain cytoplasmic immunoreceptor tyrosine-based inhibitory motifs
XX (ITIMs), whilst other LIR family members lack ITIMs. By analogy with the
XX structure and function of known MHC Class II receptor molecules, LIRs
XX having ITIMs are inhibitory receptors mediating negative signalling,
XX whilst those lacking ITIMs are activatory receptors. Failure of a
XX receptor that mediates negative signalling could result in autoimmune
XX disease. Thus, the function of a receptor mediating negative signalling
XX could result in suppressed immune function. They are also useful to
XX produce probes for detecting LIR nucleic acids or isolating LIR DNA from
XX other species.

50 Sequence 2194 BP; 479 A; 729 C; 597 G; 389 T; 0 other:

alignment_scores:

Quality: 90.50 Length: 138
Ratio: 1.560 Gaps: 6
Percent Similarity: 42.029 Percent Identity: 28.986

alignment_block:

US-09-471-276-831 x AAV69338 ..

Align seg 1/1 to: AAV69338 from: 1 to: 2194

```

1 MetSerMetLeuValValPheLeuLeuThrProGlyValThrProGlyPro 17
173 AG.....AlaGlnGlu..... 174
51 GlyProProGlyAspSerArgLeuProAlaValGlnGluTrpGly..... 65
169 .....CCAGGC.....TCTGTATACGCTGGCGGAGCC 197
66 .....AlaGlnGlu..... 68
198 CGTGACCATCTGTCGTGACGGAGGAGCCCTGAGAGCCGAGAGTACCACTG 247
69 .....ProValHisLeuAsp.....SerPro 75
248 ATAAAGAGGAGAGCCAGAGCCCTTGGAGACAAATACCCCACTGGAACCC 297
76 AlaIleLysHisGlnPheLeuLeuThrGlyAspThrGln.....G1 89
298 AAGAACAGAGGCCAGATTCCTCATCCATCCATGACAGACAGCCAGTCGAG 347

```

```

89 yAcgTgYArGcYsAqSgSerGlyLeuSerThrGlyTrp***GlnLeuSerL 106
|||||
348 GAGATACCGCTGCTCCACCTATTACACAGCTCTGACAGCTGCTCAGACCCAGCC 397
106 ySLeulEuGlnLeu 110
|||||
398 ACCCTGCGAGCTC 411

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NM2000.DNT:AAA54601
seq_documentation_block:
ID AAA54601 standard: cDNA: 2194 BP.
XX
XX
AC AAA54601:
XX
XX 11-APR-2001 (first entry)
XX
XX Leukocyte immunoglobulin like receptor coding sequence (pml17).
XX
XX Leukocyte immunoglobulin like receptor. LIR: gene therapy;
XX autoimmunity; autoimmune disorders; immune system; human; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 61..1062
XX /feature:
XX /product: Leukocyte immunoglobulin like receptor
XX
XX W0200068383-A2.
XX
XX 16-NOV-2000.
XX
XX 12-MAY-2000: 2000MO-US13228.
XX
XX 12-MAY-1999: 990US-0310463.
XX
XX (IMMEX ) IMMUNEX CORP.
XX
XX Cosman DJ, Anderson DM, Borges L;
XX WPI: 2000-667645/67.
XX P-PSDB: AAB04177.
XX
XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for
XX treating autoimmune diseases and disease states with suppressed immune
XX function
XX
XX Example 9: Page 92-95; 117pp: English.
XX
XX Host cells transformed with a vector which are capable of
XX expressing a Leukocyte immunoglobulin like receptor (LIR), can be
XX used to produce the LIR polypeptide. LIR coding sequences may be
XX used in the gene therapy of disorders mediated directly or
XX indirectly by defective or insufficient amounts of any of the LIR
XX polypeptides. The LIR polypeptides can be used to treat autoimmune
XX diseases and disease states with suppressed immune function.
XX
XX Sequence 2194 BP: 479 A; 729 C; 597 G; 389 T; 0 other:
50

```

```

|||||
67 ATGACCGCCCGCCCTCCACAGCCCTGCTGCTGCTGAGCTGAGTGGAGCC 116
17 ovalThrGlnAlaAla11pHetArgLInThrGln***SerLeuTPALAG 34
117 CAGAGCCCGCGTGACAGCAGAGGCGCTTCCGCAAGCCAGCCCTGSGGTG 166
34 LuserGlnLHis***LeuysThrLeuAlGlnGysAspAlaAspValPro 50
167 AG.....
51 GLPProGlyAspSerAsgLeuProAlaWValGlnLInThrGly ..... 65
|||||
169 .....CCAGGC.....TCTGTGATCAGCTGGGGAGAGCC 197
66 .....AlaGlnGlu..... 68
198 CGTGACCATCTGCTGTCAGAGGAGCGCTGAGGCCAGAGTACCAACATG 247
69 .....ProValHisLeuAsp.....SerPro 75
248 ATGAAGACGCAAGCCCAAGCCCTTGAGACAAAGAAATACACCACTGCAACCC 297
76 Ala1LeuHisLsGlnPheLeuLeuThrGlyAspTrpGln.....G1 89
296 AAGACAAAGCCAGATTCCTCATTCCTCCTCAGCATGACAGCAGCATGACG 347
89 yAcgTgYArGcYsAqSgSerGlyLeuSerThrGlyTrp***GlnLeuSerL 106
|||||
348 GAGATACCGCTGCTCCACCTATTACACAGCTCTGACAGCTGCTCAGAGCCAGCC 397
106 ySLeulEuGlnLeu 110
|||||
398 ACCCTGCGAGCTC 411

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NM1998.DNT:AAV69334
seq_documentation_block:
ID AAV69334 standard: DNA: 2446 BP.
XX
XX AAV69334:
XX
XX 09-FEB-1999 (first entry)
XX
XX Human LIR-pbm36-2 cDNA.
XX
XX LIR-pbm36-2: immunoregulator; leukocyte immunoglobulin-like receptor;
XX therapeutic; treatment; disorder; ITIM; MHC class I receptor; inhibitor;
XX cytoplasmic immunoreceptor tyrosine-based inhibitory motif; activator;
XX negative signalling; autoimmune disease; suppressor; LIR; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 171..1040
XX /tag: a
XX /product: "LIR-pbm36-2"
XX
XX W09848017-A1.
XX
XX 29-OCT-1998.
XX
XX 23-APR-1998: 98MO-US08244.
XX
XX 24-APR-1997: 97US-0842248.
XX
XX (IMMEX ) IMMUNEX CORP.
XX
XX Cosman DJ;
XX WPI: 1998-609990/51.
XX P-PSDB: AAN82548.
XX

```


Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful, e.g. for treating autoimmune diseases or disease states associated with suppressed immune function

Example 9: Page 71-73; 117pp: English.

This sequence encodes a novel leukocyte immunoglobulin-like receptor (LIR) polypeptide LIR-pbm36-2. This sequence can be administered therapeutically to treat disorders associated with insufficient/defective amounts of LIR polypeptide. LIR-p302 and certain other LIR family members contain cytoplasmic immunoreceptor tyrosine-based inhibitory motifs (ITIMs), whilst other LIR family members lack ITIMs. By analogy with the structure and function of known MHC Class I receptor molecules, LIRs which lack ITIMs are predicted to be activating receptors. Failure of a receptor that mediates negative signalling could result in autoimmune diseases, whilst failure of a receptor mediating activatory signalling could result in suppressed immune function. They are also useful to produce probes for detecting LIR nucleic acids or isolating LIR DNA from other species.

Sequence 2446 BP; 611 A; 692 C; 557 G; 566 T; 0 other:

Alignment_scores:
Quality: 89.50 Length: 121
Ratio: 1.377 Gaps: 7
Percent Similarity: 53.719 Percent Identity: 33.884

alignment_block:

US-09-471-276-831 x AMV69334 ...

Align seq 1/1 to: AMV69334 from: 1 to: 2446

```

1 MetSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpGlyP 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
171 ATGACCCCGCTCCACGCTCCATCTCTGTCTGTGGCTGACGTCTGGGCC 220
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 oValThrGluAlaIaIaIlePheTrpGluThrGln**SerLeuTrpAla 34
  | |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
221 CCGACCCACGCTGACGACGACCCCTCCACCCACGACACCTCTGGCGTG 270
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 IuSerGluHis**LeuLysThrLeuGly..GlnGysAspAlaAspValP 50
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
271 AGCCAGGCTCTGTGATC...ACCCAGGAGATCTCCGCTGCTCTGTCTG 317
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 roGlyProProGlyAspSerArgLeuProAlaVal..GlnGlu..... 63
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 CAGGCGATCTCTGGAGCCAGGAGTACCTCTGTTATGAGAGAAAGAAAC 367
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 .....TTPGlyAlaGlnGluProValHisLeuAspSerProAlaIleL 78
368 AGACACCTGAGTATACAGCATCCACAGCAAT.....GTAAAGC 408
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
78 yshISGlnPhe.....LeuLeuThrGlyAspThrGlnGlyArgTrp 91
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
409 AGGCGCATCTCCCATCCCTCCATACCTCGGAGACACCGCGGCGCTAT 458
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
92 ArgGlyAspSerGlyLeuSerThr..GlyTrp**GlnLeuSerLysLe 107
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 CGGTGTTTTCACGTTAGCCACACTGACAGCTGCTGTACAGCCAGTACC 508
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 uLeuGluLeu 110
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
508 CTRTGAGCTG 518
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
seq_name: /SID52/gcysdata/geneseq/geneseq/NR2000.DAT:AAA54597
seq_documentation_block:
ID AAA54597 standard: CDNA; 2446 BP.
XX
AC
AA54597:
XX
```

DT 11-APR-2001 (first entry)

DE Leukocyte immunoglobulin like receptor coding sequence (pbm36-2).

XX Leukocyte immunoglobulin like receptor; LIR: gene therapy; LIR
XX autoimmunity; autoimmune disorders; immune system; human; ds.

CS Homo sapiens.

PH Key Location/Qualifiers

FT CDS /tag= a 171..1040

FT /product= Leukocyte immunoglobulin like receptor

PN W0200068383-A2.

PD 16-NOV-2000.

PF 12-MAY-2000; 2000MO-US13228.

XX 12-MAY-1999; 9905-0310463.

PA (IMW) IMMUNEX CORP.

PT Cosman DJ, Anderson DM, Borges L;

XX WP1: 2000-687645/67.

XX P-PSDB: AAB04173.

XX Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for

PT treating autoimmune diseases and disease states with suppressed immune

XX function

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```

69 .....Prox1h1stLeuAsp.....:SepPro 75
231 ATAAAGAGGAGACCCGAGGCCCTTGAGCGAGAAATACCCACTGAGAACCC 280
76 A1a1a1y1s1s1sc1p1e1n1e1u1e1r1c1y1a1s1p1t1h1a1.....:C1 89
281 AAGACAAAGCCAGCAATTCCTCCATCCATCCATCAAGCAAGACCAACATCGGCG 330
89 y1a1g1T1a1c1g1y1a1s1g1e1c1y1e1u1s1e1r1t1c1y1T1P1**G1n1Leu1Se1L1 106
331 GAGATACCGCTCGGCACTATTCACAGCTCTGCAAGCGCTGTCAGACCGACCGACG 380
106 y1a1e1u1e1n1g1l1e1u1d1e1 110
381 ACCCTCGGAACTG 394

seq_name: /SID52/3/gcodb/ty/geneseq/geneseq/NA1998.DAT:AAV63335
seq_documentation_block:
ID AAV6335 standard; DNA: 1910 BP.
XX
XX
XX AAV6335:
XX
XX
XX 09-FEB-1999 (first entry)
XX
XX
XX Human LIR-pbm36-4 cDNA.
XX
XX
XX LIR-pbm36-4: immunoregulator; leukocyte immunoglobulin-like receptor;
XX therapeutic; treatment; disorder; ITIM; MHC class II receptor; inhibitor;
XX cytoplasmic immunoreceptor; cytosine-based inhibitory motif; activator;
XX negative signalling; autoimmune disease; suppressor; LIR; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 183..1652
XX FT CDS
XX ET /*seq= a
XX FT /product= "LIR-pbm36-4"
XX
XX MO9648017-A1.
XX
XX
XX 23-APR-1998: 98W0-US08244.
XX
XX 24-APR-1997: 97US-0842248.
XX
XX (IMM) IMMDEX CORP.
XX
XX
XX Cosman DJ:
XX
XX WPI: 1998-609990/51.
XX P-PSDB: AAW62549.
XX
XX Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful,
XX e.g. for treating autoimmune diseases or disease states associated
XX with suppressed immune function
XX
XX Example 9: Page 74-77, 112pp; English.
XX
XX
XX This sequence encodes a novel leukocyte immunoglobulin-like receptor
XX (LIR), polypeptide LIR-pbm36-4. This sequence can be administered
XX therapeutically to treat disorders associated with insufficient/defective
XX amounts of LIR polypeptide. LIR-p3c2 and certain other LIR family members
XX contain cytoplasmic immunoreceptor tyrosine-based inhibitory motifs
XX (ITIMs), which anchor LIR family members to the cell surface. LIRs
XX having ITIMs are inhibitory receptors mediating negative signalling,
XX whilst those lacking ITIMs are activatory receptors. Failure of a
XX receptor that mediates negative signalling could result in autoimmune
XX diseases, whilst failure of a receptor mediating activatory signalling
XX could result in suppressed immune function. They are also useful to

```

CC produce probes for detecting LIR nucleic acids or isolating LIR DNA from
CC other species.

XX Sequence 1910 BP; 408 A; 588 C; 509 G; 405 T; 0 other:

alignment_scores:

Quality: 84.50 Length: 121
Ratio: 1.320 Gaps: 7
Percent Similarity: 52.893 Percent Identity: 33.884

alignment_block:

US-09-471-276-831 x AAV69335 ..

Align seg 1/1 to: AAV69335 from: 1 to: 1910

```

1 MelSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpPolypC 17
||||: |||||: |||||: |||||: |||||
183 ATGACCCCGCATGTCACAGTCCTCATCTGTCTGACGCTGACGCTGCGGCC 232
17 cvalThGtAlaAla11ePheTygIuThGin***SerLeuTrpAlaG 34
||||: |||||: |||||: |||||: |||||
233 CCGGACCCAGTCGACGAGGACCGCTCCGACAGCCACACTGCTGGCTG 282
34 IuSerGluHis**LeuTyThrLeuLy..GIncyAspAlaAspValP 50
||| ||| ||| ||| |||
283 AGCCAGGCTCTGTATC...ACCAGGGAGATCCCTGACACCTCTGTCT 329
50 roGlyProProGlyAspSerArgLeuProAlaVal..GInGlu..... 63
||||: |||||: |||||: |||||: |||||
330 CAGGAGATCCTGGAGACCCAGGACGACGCTGTCTGTATGAGAAAGAAAC 379
64 .....TrpGlyAlaGInGluProValHisLeuAspSerProAla11eL 78
380 AGCACCCTGGATTCACAGCATCCACAGAGATTT.....GTGAGAG 420
78 yThGInGlnPhe.....LeuLeuThGlyAspThrGInGlyArgTrp 91
|| ||||| ||||| ||||| |||||
421 AGGCGCAGTTCGCCATCCATCCATCCGCTGCGACACACAGGCGCGAT 470
92 ArgCyshArgSerGlyLeuSerThr...GlyTrp**GlnLeuSerLyLe 107
||||| ||||| ||||| ||||| |||||
471 CGCTGTCTTTCAGGCTAGCCACACCTGACGAGCTGTCAAGCCAGTACC 520
107 uLeuGlnLeu 110
||||| |||||
521 CCTGGACACTG 530

```

seq_name: /SID52/gcgnata/geneseq/geneseq/NM2000.DAT:AAAS4598

seq_documentation_block:

ID: AAAS4598 standard; CDNA; 1910 BP.

AC AAAS4598;

XX 11-Apr-2001 (first entry)

XX Leukocyte Immunoglobulin like receptor coding sequence (pbm36-4).

XX Leukocyte Immunoglobulin like receptor; LIR; gene therapy;

XX autoimmunity; autoimmune disorders; immune system; human; ds.

XX Homo sapiens.

OS Key Location/Qualifiers

XX CDS 1910-692 /Join: 692

XX FT /Product: Leukocyte Immunoglobulin like receptor

XX MO200068383-A2.

XX 16-NOV-2000.

PF 12-MAY-2000: 2000MO-US13228.

XX 12-MAY-1999: 9905-0310463.

XX (IMMUNEX CORP.

XX Cosman DJ, Anderson DM, Borges L;

XX WPI: 2000-687645/67.

XX P-PSDB: AA004174.

XX Leukocyte Immunoglobulin-like receptor (LIR) nucleic acids, useful for

XX treating autoimmune diseases and disease states with suppressed immune

XX function

XX Example 9; Page 80-83, 117pp: English.

XX Host cells transformed with a vector which are capable of

XX expressing a leukocyte immunoglobulin like receptor (LIR), can be

XX used to produce the LIR polypeptide. LIR coding sequences may be

XX used in the gene therapy of disorders mediated directly or

XX indirectly by the LIR polypeptides. The LIR polypeptides can be used to treat autoimmune

XX diseases and disease states with suppressed immune function.

XX sequence 1910 BP; 408 A; 588 C; 509 G; 405 T; 0 other:

alignment_scores:
Quality: 84.50 Length: 121
Ratio: 1.320 Gaps: 7
Percent Similarity: 52.893 Percent Identity: 33.884

alignment_block:

US-09-471-276-831 x AAAS4598 ..

Align seg 1/1 to: AAAS4598 from: 1 to: 1910

```

1 MelSerMetLeuValValPheLeuLeuLeuTrpGlyValThrTrpPolypC 17
||||: |||||: |||||: |||||: |||||
183 ATGACCCCGCATGTCACAGTCCTCATCTGTCTGACGCTGACGCTGCGGCC 232
17 cvalThGtAlaAla11ePheTygIuThGin***SerLeuTrpAlaG 34
||||: |||||: |||||: |||||: |||||
233 CCGGACCCAGTCGACGAGGACCGCTCCGACAGCCACACTGCTGGCTG 282
34 IuSerGluHis**LeuTyThrLeuLy..GIncyAspAlaAspValP 50
||| ||| ||| ||| |||
283 AGCCAGGCTCTGTATC...ACCAGGGAGATCCCTGACACCTCTGTCTGT 329
50 roGlyProProGlyAspSerArgLeuProAlaVal..GInGlu..... 63
||||: |||||: |||||: |||||: |||||
330 CAGGAGATCCTGGAGACCCAGGACGACGCTGTCTGTATGAGAAAGAAAC 379
64 .....TrpGlyAlaGInGluProValHisLeuAspSerProAla11eL 78
380 AGCACCCTGGATTCACAGCATCCACAGAGATTT.....GTGAGAG 420
78 yThGInGlnPhe.....LeuLeuThGlyAspThrGInGlyArgTrp 91
|| ||||| ||||| ||||| |||||
421 AGGCGCAGTTCGCCATCCATCCATCCGCTGCGACACACAGGCGCGAT 470
92 ArgCyshArgSerGlyLeuSerThr...GlyTrp**GlnLeuSerLyLe 107
||||| ||||| ||||| ||||| |||||
471 CGCTGTCTTTCAGGCTAGCCACACCTGACGAGCTGTCAAGCCAGTACC 520
107 uLeuGlnLeu 110
||||| |||||
521 CCTGGACACTG 530

```

seq_name: /SID52/gcgnata/geneseq/geneseq/NM1998.DAT:AAV69339

seq_documentation_block:

ID	AAV69339	standard: DNA: 2061 bp.
AC	AAV69339:	
XX	09-FEB-1999	(first entry)
XX		
XX	Human LIR-phmnew cDNA.	
XX	LIR-phmnew: immunoregulator; leukocyte immunoglobulin-like receptor;	
XX	therapeutic; treatment; disorder; TIM; MHC class I receptor; inhibitor;	
XX	cytoplasmic immunoreceptor tyrosine-based inhibitory motif; activator;	
XX	negative signalling; autoimmune disease; suppressor; LIR: ss.	
XX		
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
PH	CD5	67..1839
FT		/*tag= a
FT		/product= "LIR-phmnew"
XX		
XX	MO9848017-A1.	
XX	29-OCT-1998.	
XX		
XX	23-APR-1998:	98MO-US08244.
XX	24-APR-1997:	97US-0842248.
XX	(IMMY) IMMUNE CORP.	
XX		
XX	Cosman DJ:	
PI		
XX	WPI: 1998-609990/51.	
DR	P-PSDB: AANR2553.	
XX		
XX	Leukocyte immunoglobulin-like receptor, LIR, polypeptides - useful,	
XX	with suppressed immune function	
XX		
XX	Example 92: Page 92-94; 112pp: English.	
XX		
XX	This sequence encodes a novel leukocyte immunoglobulin-like receptor	
CC	(LIR) polypeptide LIR-phmnew. This sequence can be administered	
CC	therapeutically to treat disorders associated with insufficient/defective	
CC	cytoplasmic immunoreceptor tyrosine-based inhibitory motifs	
CC	contain cytoplasmic immunoreceptor tyrosine-based inhibitory motifs	
CC	(ITIMs), whilst other LIR family members lack ITIMs. By analogy with the	
CC	structure and function of known MHC Class I receptor molecules, LIRs	
CC	having ITIMs are inhibitory receptors mediating negative signalling,	
CC	whilst those lacking ITIMs are activatory receptors. Failure of a	
CC	receptor that mediates negative signalling could result in autoimmune	
CC	diseases, whilst failure of a receptor mediating activatory signalling	
CC	could produce probes for detecting LIR nucleic acids or isolating LIR DNA from	
CC	other species.	
XX		
XX		
XX	Sequence 2061 bp; 431 A; 684 G; 558 C; 388 T; 0 other:	
XX		
XX	Alignment_scores:	
XX	Quality: 84.50	Length: 139
XX	Ratio: 1.207	Gaps: 9
XX	Percent Similarity: 50.360	Percent Identity: 30.935
XX		
XX	alignment_block:	
XX	US-05-471-276-831 x AAV69339	
XX		
XX	Align seg 1/7 to: AAV69339 from: 1 to: 2061	
XX		
XX	1 MerserMGLauValValPheIeuLeuLeuTPSPSjYValThiTrcGlyPT 17	
XX	: : : : : : : : :	
XX	67 ATGACCTCACCCTCTCACTGCTGATTTGGCTCTGGGCTGATGATGAGTGAGCC 116	

seq. documentation block:	seq. name: /SI052/9cde4c7c/geneseq/geneseq/MA2000. DAT: AAA54602
1 ID AAA54602 standard; cDNA: 2061 BP.	
2 AC	
3 CC	
4 AAA54602:	
5	
6 * 11-Apr-2001 (11st entry)	
7 XX	
8 Leukocyte immunoglobulin like receptor coding sequence (phmew).	
9 DE	
10 Leukocyte immunoglobulin like receptor: LIR, gene therapy;	
11 autoimmunity; autoimmune disorders; immune system; human; ds.	
12 XX	
13 OS Homo sapiens.	
14 XX	
15	
16 Key Location/Qualifiers	
17 FT CDS 67..1839	
18 FT /tag= a	
19 FT /product= leukocyte immunoglobulin like receptor	
20 PN MO200068383-A2.	
21 XX	
22 16-NOV-2000.	
23 PD	
24 12-MAY-2000; 2000NO-US13228.	
25 PF	
26 12-MAY-1999; 99US-0310463.	
27 PF	
28 (TMV) IMMUNEX CORP.	
29 PA	
30 Cosman DJ, Anderson DM, Borges L;	
31 PI	
32 WPI: 2000-687645/67.	
33 DR	
34 P-PSDB: AA04178.	
35 Leukocyte immunoglobulin-like receptor (LIR) nucleic acids, useful for	
36 treating autoimmune diseases and disease states with suppressed immune	
37 function	
38 XX	
39	
40 Disclosure: Page 97-100; 117pp; English.	
41 XX	

PA (IMM) IMMUNEX CORP

DT	23-SEP-1998	(first entry)
XX		

DT 23-SEP-1998 (first entry)

[illegible]

PD 22-FEB-2000

OS Homo sapiens.

```

-9-471-276-831 x AA55010
433 ATGGCTGTAATTTCTGTCGGGATGACCTGATCGACGGCGGAA 482
434 AGCGTCTCTCCGGGCMCAACCTCTTCTGCACAGCGGGAGCTGGACACTCA 533
20 ualalallelePheTyrGluPheTrpIle**Serine.....TTPAlaGlu 35
435 ualalallelePheTyrGluPheTrpIleTTPGlyProValTrpIle 20
436 LenvyYAlPhePheLeuLeuLeuPheTyrGlyValTrpIleTTPGlyProValTrpIle 20
437 ATGGCTGTAATTTCTGTCGGGATGACCTGATCGACGGCGGAA 482

```

XX Claim 7, page 21-22: 30pp: Japanese.
 CC The present invention describes new proteins for inhibiting
 CC the development of atherosclerosis and other diseases. The phosphatase
 CC SHP-1, SHP-2 or inositol-5-phosphatase SHIP by phosphorylating tyrosine-
 CC The present invention also describes an immunoreceptor comprising one
 CC of the above proteins, and DNA coding the above proteins. The new proteins
 CC can be used as an immunomodulatory agent. The present sequence encodes
 CC a protein from the present invention.
 XX
 XX Sequence 1906 BP; 518 A; 546 C; 419 G; 423 T; 0 other:
 alignment_scores:
 Quality: 75.50 Length: 114
 Ratio: 1.302 Gaps: 5
 Percent Similarity: 50.877 Percent Identity: 28.947
 alignment_block:
 US-09/471-27/6-831 x AAX88978 ..
 Align seq 1/1 to: AAX88978 from: 1 to: 1906
 6 ValPheMetLeuValIlePheValThrTTPGlyProValIleThcGluAla.. 21
 ||| |||||::: |||
 |||||::: |||

```

643 GTGAGAGCTCGTGCAGTAACTCCGAAACCAACCATCAGAGCTCA 692
22 .....AlaIlePheTyrgIuThrGln***SerLeuProAlaIaGlu 35
693 ACCAGAGATCTGAGTGGCTCCCAAAAGCAACATGACATCTGGTGCAG 741
35 ecrluHLS***LeuIysThrLeuGlcIcngAspAlaIaAspVal..... 49
742 .....CGCAACCTCGATCGACAGATATATTTT 768
50 ProGlyProProGlyAspSerArgLeuProAlaValGlnGluTrpAlaI 66
769 CTCGATTAATGAGGAGAGCAAAACCAACAGCAACAGC..... 810
66 ncIuGluProValHisLeuAspSerProAlaIleIysThrAlaGluPheLeu 83
811 .....CTACACAGACCTGGGACACAGGAGCTGCTTCA 844
83 euThiGlyAspThrGln.....GlyArgTyArgGlyAspGrcIy 96
845 TCCTCTCTGTCGACACATGCAAGGCAATGCTGCTTATGTTAC 894
97 LeuSerThcIcIyTrp***GlnLeuSerTysLeuLeuGlu 110
895 ACCTGACCTGCTGTCACACCCGACGTGACACCCGACGCTG 936
seq_name: /SID52/gcgdata/geneseq/geneseq/NM2000.DMT:AAF18310
seq_documentation_block:
ID AAF18310 standard: DNA: 2373 BP.
AC
AC AAF18310:
AT
AT 14-MAR-2001 (first entry)
DE
DE Lung cancer associated polynucleotide sequence SEQ ID 329.
XX
XX Human: lung cancer associated protein; neuroprotective; cytostatic;
XX cardioprotective; immunomodulatory; muscular active; vulnerary;
XX gastrointestinal; nephrotoxic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.
OS
OS Homo sapiens.
XX
XX W0200055180-A2.
PN
PN 21-SEP-2000.
PD
PD 08-MAR-2000: 2000MO-US05918.
XX
XX 12-MAR-1999: 59US-0124270.
PR
PR (HUMA-) HUMAN GENOME SCI INC.
PA
PA (ROSE-) ROSEN C. A.
XX
XX Ruben SM:
XX
XX WPI: 2000-587514/55.
DR
DR P-PSDB: AAB58434.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as lung cancer -
XX
XX Claim 1: Page 786-787: 1425pp: English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX associated proteins and polynucleotide sequences, their agonists, and
XX antagonists may have neuroprotective, cytostatic, cardioprotective,
XX immunomodulatory, muscular active general, vulnerary, gastrointestinal
XX general, nephrotoxic, antiinfective, gynecological, or antibacterial

```

CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC lung cancer, immunomodulation, neuroprotection, wound healing and proliferative
 CC gastrointestinal pulmonary cardiovascular, renal and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58949 are used in the course of the invention for the
 CC identification and characterization of the polynucleotide and protein
 CC sequences.

Sequence 2373 BP: 456 A: 730 C: 669 G: 514 T: 4 Other:

Alignment_scores:

Quality:	75.50	Length:	90
Ratio:	1.641	Gaps:	5
Percent Similarity:	51.111	Percent Identity:	27.778

Alignment_block:

us-09-471-276-831 x AAF18310 ..
 Align seg 1/1 to: AAF18310 from: 1 to: 2373

```

4 LeuValValPheLeuLeuTrpGlyValThrTrpGlyProValThrG 20
.....
733 ATGGCTGTAATTTCTGTGGGATGATGATGATGATGATGATGATGATG 782
20 uAlaIaIlePheTyrgIuThrGln***SerLeu.....TrpAlaGlu 35
.....
783 AGCTGCTCTGTCGACACACCTCTCTGACAGAGCGGGGAGCTGCAC 832
35 ecrluHLS***LeuIysThrLeu.....CysAspAlaIaAspValr 42
833 CACACATCTGCTACACCTCACTTTCGTTGCGGAGACATCATCTGTT 882
43 .....GlyGln.....CysAspAlaIaAspValr 50
883 ATGAGAGATCTGAGAGCTGCTGTTATCGGCTCCACAGCGAGCTGC 932
50 ocIyProProGlyAspSerArgLeuProAlaValGlnGluTrpGly... 65
.....
933 CTCACACAGGAGATTCGGCTTTCCTGTCGAGATGATGATGATGATG 982
66 AlaGlnGluProValHis 71
.....
983 CTCACGCAATGATTCAT 1000

```

Page 1

no sapiens
no sapiens
no sapiens
no sapiens
no sapiens

[illegible]

00001
SPLING IN
COE Lomi ;

00001
NC IN
00001

Ratio: 1.612 Gaps: 6
Percent Similarity: 55.833 Percent Identity: 34.167
alignment_block:
US-09-471-276-831 x AF324830 ..

Align seg 1/1 to: AF324830 from: 1 to: 1365

1 MetSerMetLeuValPheLeuLeuThrProGlyValThrThrProLysr 17
175 GTGAGCCGCCCTCCATGCTGCTGCTCCCTCCGCTGAGCTGGCC 224
17 ova1ThgGluAlaAlaIlePheArgGluThrGln***SerLeuThrAlaG 34
175
228 CAGAGCCAGCTGCGAGCAGACACTCCAAAGCCACCTCTGGCGCTG 274
34 InSerGluHis**LeuLysThrLeuLysGln..CysAspAlaAspValP 50
275 AGCCAGGCTCTGTATGCT...AGCGGGGAGACCTCTGACACCTCGGT 321
50 nclGlyProProGlyAspSerArgLysLeuProAlaValGlnGlu 63
322 CAGGAGACCTCGAGCCCGAGCATACCTCTGTTTAAAGAGAGACCC 371
64TTPGlyAlaGlnGluProValHisLeuAspSerProAlaIle 78
372 AGAGCCCTGGAGACACAGAACCCACTG.....GAGCCAGAGACA 412
78 yst1aGlnPheLeuLeuThrGlyAspThrGln.....GlyArgThr 91
413 AGCCGATTCTCCATCCATCCATCCAGACGACGACGACGACGACGAC 462
92 ATGCGAATGergArgLysSerThrGlyTTP**GlnLeuSerLysLeu 108
463 GCGCTTACTACTACACCCCTGAGCGAGCTGTGAGACCGAGACCCCT 512
108 uGluLeu 110
513 GAGACTG 519
seq_name: gbl_pr:HS091925
seq_documentation_block: 1603 bp mRNA 03-SEP-1997
LOCUS HS091925
DEFINITION Human clone HM18 monocytic inhibitory receptor precursor mRNA,
complete cds.
ACCESSION U91925
VERSION U91925.1 GI:2351798
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Alm,J.P., Nwanikwo,C. and Austen,K.F.
TITLE Molecular identification of a novel family of human immunoglobulin
superfamily members that possess immunoreceptor tyrosine-based
receptor motifs and homology to the mouse gp96b1 inhibitory
receptor
JOURNAL Unpublished
AUTHORS Alm,J.P.
TITLE 2 (bases 1 to 1603)
REFERENCE Submitted (04-MAR-1997) Medicine, Harvard Medical School and
JOURNAL Brigham MA 02146 USA
Boston, Location/Qualifiers
1..1603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map_19q13.4"
/cell_type="monocyte"

5'UTR
CDS
/clone="HM18"
/c1..45
/cdon_start=3
/cdon_end=1392
/protein_id="AA68665.1"
/protein_id="AA68665.1"
/db_xref="GI:2351799"
/translation="MTTPALALGLSLGLOPRTIMAGCPILKPTLMAKPSVLSMGNS
VTWCOTLEAREHYHLKESHPANDRONPLPKANKAPSP1PMTEDYAGVACYS
PWGSSPSIPLEIVTGAASVKPRTLSALPSLVTSQKSVTLGCSHPMTGLIKVA
APRPLILSRBKAGQIDQAPRPMISVYSNQCITNCSLSHGSHYLLSHSHLEILVS
SELSRPSKSKSAGKQVDEQDQVPSPPKAPRPMISVYSNQCITNCSLSHGSHY
NTPRSGVNDQSPHIDRQAVNYAKVHSRBRBAASPSPLSGERLDTDRQAC
EDRQDPTEAASRAGQVYTAQHLSTPLKQKATPPSSQECASPAEVSATLAIH"
45..114
s1g_peptide
mal_peptide
misc_feature
115..1389
/product="monocyte inhibitory receptor"
misc_feature
115..822
/note="encodes cytoplasmic domain"
misc_feature
822..1389
/note="encodes transmembrane domain"
misc_feature
886..1389
/note="encodes cytoplasmic domain"
misc_feature
1117..1134
/note="encodes ITIM motif"
misc_feature
1273..1290
/note="encodes ITIM motif"
3'UTR
polyA_signal
polyA_site
1378..1583
1603
/note="14 A nucleotides"
BASE COUNT 384 a 517 c 427 g 275 t
ORIGIN
alignment_scores:
Quality: 103.50 Length: 121
Ratio: 1.617 Gaps: 7
Percent Similarity: 52.893 Percent Identity: 36.364
alignment_block:
US-09-471-276-831 x HS091925 ..
Align seg 1/1 to: HS091925 from: 1 to: 1603
2 SerMetLeuValPhe.....LeuLeuLeuThrProGlyValThrThrProL 16
43 GTGAGCCGCCCTCCATGCTGCTGCTCCCTCCGCTGAGCTGGCC 92
16 yova1ThgGluAlaAlaIlePheArgGluThrGln***SerLeuThrAlaG 33
93
93 CAGAGCCAGCTGCGAGCAGACACTCCCAAGCCACCTCTGGCG 142
33 lAGLysSerGluHis**LeuLysThrLeuLysGln..CysAspAlaAsp 48
143 GTGAGCCGCCCTCCATGCTGCTGCTCCCTCCGCTGAGCTGGCC 188
49 ValProGlyProProGlyAspSerArgLysLeuProAlaValGlnGlu 63
189 GTGAGCCAGCTGCGAGCAGACACTCCCAAGCCACCTCTGGCG 238
64TTPGlyAlaGlnGluProValHisLeuAspSerProAla 76
239 GCGGAGACCTCGAGACACAGAACCCACTG.....GAGCCAGAGACA 279
77 lLysHisGlnPhe.....LeuLeuThrGlyAspThrGlnGlyVal 90
280 AACAGAGCCAGATTCATCCATCCATCCAGACGACGACGACGACGAC 329
90 gTYArgGysArgSerGlyLysSerThrGlyTTP**GlnLeuSerLysL 107
330 ATGCGCTTACTACTACACCCCTGAGCGAGCTGTGAGACCGAGACCC 379

```

107 eueugluen 110
|||||
380 CCCTGGAGCTG 390
seq_name: gb_pat:AX046954
seq_documentation_block:
LOCUS AX046954 1625 bp DNA PAT 15-DEC-2000
DEFINITION Sequence 17 from Patent M0060383.
ACCESSION AX046954
VERSION AX046954.1 GI:11876384
SOURCE human
ORGANISM Homo sapiens
Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1625)
Cosman,D.J., Anderson,D.M. and Borges,L.
Family of Immunoregulators and Derived Leukocyte Immunoglobulin-like
Peptide: NO 0068383-A 17 16-NOV-2000;
JOURNAL IMMUNEX CORPORATION (US)
FEATURES
source location/Qualifiers
1..1625
/organism="Homo sapiens"
/db_xref="taxon:9606"
30..1376
/codon_start=1
/protein_id="CAC18933.1"
/db_xref="GI:11876385"
/translation="MIPPTALACGLSLQPTTHMOAGPLPKPTLMANGSVLSNGNS
VTWCGQTLFARVYLDKESPAHNDRONLPFNKNARSPISPTEDYAGRYCYKRS
PAGNSQSPLELVATGASKPTLSALSPVLTSGSVTLTLCOSPMOTFLILERA
ATGCRPSGADGACGAPPSVYSHGCTTCTCSHGGSHITLASHSPVLELTVS
GATGCRPSGADGACGAPPSVYSHGCTTCTCSHGGSHITLASHSPVLELTVS
FLILQHWCKHRTLAQRADQRPDPCGAAPPRDGLQRRSPVAALVGGNCANVX
NTPQEDGVMDTRPSHDIDPOAVYTAAYVNSRRRRMAKSPSPISGERTDTKQAR
EDROMDTPAASEADQVYTAAYVTAHSPFLRQKATPPPSQDASPAEPSVATLAIH"
BASE COUNT 399 a 523 c 425 g 278 t
ORIGIN
alignment_scores:
Quality: 103.50 Length: 121
Ratio: 1.617 Gaps: 7
Percent Similarity: 52.893 Percent Identity: 36.364
alignment_block:
US-09-471-276-831 x AX046954 ..
Align seg 1/1 to: AX046954 from: 1 to: 1625
2 SerMetLeuValValPhe.....LeuLeuLeuTrpGlyValThrTrpG1 16
|||||
27 GCCTGATGCCACCTTCACGGCTCTCTCTCTCTCTCTCTCTGATCTGCG 76
|||||
16 yProValThrGluAlaAlaIlePheTrpGluTrpGln***SerLeuTrpA 33
|||||
77 CCCAGAGCCACATGCACAGAGGGCCCCCTCCCAACACCCACCTCGAG 126
|||||
33 IagLysGluIleHis**LeuLysThrLeuGlyGln...CysAspAlaSer 48
|||||
127 CTGAGCGAGGCTCTGTGATCAG....CTGGGGGACCTCTGTACCATCTG 172
|||||
49 ValProGlyProProGlyAspSerArgLeuProAla ValGlnGln... 63
|||||
173 GTTCACAGGGAACCTGTGAGAGCTGGAGACCTCTGATCTGATGATAGAA 222
|||||
64 .....TrpGlyAlaIagLagLagProValHisLeuAspSerProAla 76
|||||
223 GCCACAGACCCCTGGGGACAGACAGAACCCACATG.....GACCCCAAG 263

```

```

77 IleLysHisLeuPhe.....LeuLeuThrGlyAspThrGlnGlyAc 90
|||||
264 MACAGAGCCAGATTCATCCATCCATCCATCCATCCATCCATCCATCCAGGAG 313
|||||
90 yTPARArgGln*LeuGlyLeuSerThrGlyValPhe***GluLeuSerLysL 107
|||||
314 ATACCGCTGTACTATCCAGACCGCTGTAGGCTGTCTACACACCCATGAGCC 363
|||||
107 eueugluen 110
|||||
364 CCCTGGAGACTG 374
seq_name: gb_pr:AF025532
seq_documentation_block:
LOCUS AF025532 1625 bp mRNA PRI 01-DEC-1997
DEFINITION Homo sapiens leucocyte immunoglobulin-like receptor-5 (LIR-5) mRNA,
complete cds.
ACCESSION AF025532
VERSION AF025532.1 GI:2653870
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1625)
Borges,L., Hsu,M.-L., Fanger,N., Kublin,M. and Cosman,D.
A family of human lymphoid and myeloid immunoglobulin-like
leucocyte surface receptors and to MHC class I molecules
J Immunol. (1997) In press
2 (bases 1 to 1625)
Borges,L., Hsu,M.-L., Fanger,N., Kublin,M. and Cosman,D.
Direct Submission
Submitted (18-SEP-1997) Molecular Biology, Immunex Corp., 51
University St., Seattle, WA 98101, USA
JOURNAL
ORIGIN
FEATURES
source location/Qualifiers
1..1625
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1625
/gene="LIR-5"
/gene="LIR-5"
30..1376-5*
/codon_start=1
/product="leucocyte immunoglobulin-like receptor-5"
/protein_id="AA87666.1"
/db_xref="GI:2653871"
/translation="MIPPTALACGLSLQPTTHMOAGPLPKPTLMANGSVLSNGNS
VTWCGQTLFARVYLDKESPAHNDRONLPFNKNARSPISPTEDYAGRYCYKRS
PAGNSQSPLELVATGASKPTLSALSPVLTSGSVTLTLCOSPMOTFLILERA
AHPRLSHKEMGQDALEPRPSVYSHGCTTCTCSHGGSHITLASHSPVLELTVS
FLILQHWCKHRTLAQRADQRPDPCGAAPPRDGLQRRSPVAALVGGNCANVX
NTPQEDGVMDTRPSHDIDPOAVYTAAYVNSRRRRMAKSPSPISGERTDTKQAR
EDROMDTPAASEADQVYTAAYVTAHSPFLRQKATPPPSQDASPAEPSVATLAIH"
BASE COUNT 399 a 523 c 425 g 278 t
ORIGIN
alignment_scores:
Quality: 103.50 Length: 121
Ratio: 1.617 Gaps: 7
Percent Similarity: 52.893 Percent Identity: 36.364
alignment_block:
US-09-471-276-831 x AF025532 ..
Align seg 1/1 to: AF025532 from: 1 to: 1625
2 SerMetLeuValValPhe.....LeuLeuLeuTrpGlyValThrTrpG1 16
|||||
27 GCCTGATGCCACCTTCACGGCTCTCTCTCTCTCTCTCTCTGATCTGCG 76
|||||

```



```

3'UTR
polyA_signal 1235 a 1550 c 1557 g 1012 t
ORIGIN
1081 GCATCT 1086
seq_name: gb.AF189768
seq_documentation_block:
LOCUS AF189768 8872 bp DNA PRI 27-AUG-2000
DEFINITION Homo sapiens leukocyte immunoglobulin-like receptor 5 (LIR5) gene,
complete cds.
ACCESSION AF189768
VERSION AF189768.1 GI:9930102
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Cerebralia; Homalidae; Homo.
REFERENCE
1 (bases 1 to 8872)
Mammalia; Eutheria; Primates; Cerebralia; Homalidae; Homo.
AUTHORS Liu, W.-R., Kim, J., Nwanke, C., Ashworth, L. K. and Arm, J. P.

```

```

TITLE
Genomic organization of the human leukocyte immunoglobulin-like
receptors within the leukocyte receptor complex on chromosome
19q13.4
JOURNAL
Immunogenetics 51 (8-9), 659-669 (2000)
PUBLISHED
20035285
MEDLINE
10941837
REFERENCE
2 (bases 1 to 8872)
AUTHORS
Arm, J. P.
TITLE
Direct Submission
JOURNAL
Submitted (24-SEP-1999) Medicine, Brigham and Women's Hospital, One
Jimmy Fund Way, Boston, MA 02115, USA
FEATURES
1. 8872
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4: between D19S781 and D19S268"
join(3061..3456,3933..3968,4155..4439,4587..4886,
5200..5250,5531..5581,6216..6332,6641..6716,6800..6937,
7099..7195,8036..8194,8274..8630)
/gene="LIR5"
/product="leukocyte immunoglobulin-like receptor 5"
3061..8630
/gene="LIR5"
join(3423..3456,3933..3968,4155..4439,4587..4886,
5200..5250,5531..5581,6216..6332,6641..6716,6800..6937,
7099..7195,8036..8194,8274..8420)
/note="NM18: immunoglobulin-like transcript 3"
/codon_start=1
/product="leukocyte immunoglobulin-like receptor 5"
/protein_id="AA02024.1"
/db_xref="GI:9930103"
/translation="MPTPLALGAGLSGPTFMQNGAPRPLTAAAGCSVSNKS
PVGSGPSDFLELVNGATVATSKPTSLAPLSTSGKSSISITLASHSPLELIVS
GSLGKPRPSPTSVSTACPPDPLPMTGVSFVSHGCRHMEVTLGVTVSTLTLIL
AHPLILHSEIKAGQIOADAKRWGAPVTSVHGCTVSGKSFISITLASHSPLELIVS
GSLGKPRPSPTSVSTACPPDPLPMTGVSFVSHGCRHMEVTLGVTVSTLTLIL
FLILQIMCKKNTLACROAFORPCALAEPRKDCGLORSSPADVGGSCANVAK
NTQPEGVNDTQKSPHIDPOAVTVYAKVHSRPRRREKASPSRISGFIQTDHQA*
EDRMQNTDAASPAADQVTVYAKVHSRPRRREKASPSRISGFIQTDHQA*

```


REFERENCE
1 (bases 1 to 2922)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

66 6glnclu.....Provalh1sleuasp.... 73
      |||||
26190 CAGAGAGACACAGCTGATTAAGAGGAGGAGCCAGAGCCCTTGACAGAA 26239
      |||||
74 .....SerProvalh1leu1sleu1p1heuleu1thrc1yap 86
      |||||
26240 ATAAACCACTGACACGACGACGACGATTCCTCATTCCATCCATCAG 26289
      |||||
87 Thrc1n.....GlytGlytArGcYArGserGlyleuSerThrc1 100
      |||||
26290 KACAGACACATGACAGAGATACCGCTGACATTCATACGCTCTCCAG 26339
      |||||
100 YTrp***GlnleuSerLysleu1sleu1p1heuleu1thrc1yap 111
      |||||
26340 CTGCTGACAGCCAGCAGCAGCCCTGACGCTGATGACAGCTGACAGCA 26389
      |||||
112 .....GlyProLysValLeuAlaCysSerLeu.....AlaLeu 122
      |||||
26390 CACTGCGGGGTCCACAGCCGCTGCGCTGACGAGAGGCGGTGCGCTC 26439
      |||||
123 ArgP1AlaSer 126
      |||||
26440 TCACGGCGGCTCT 26451
      |||||
seq_name: gb_pf:AF041262
seq_documentation_block: 1395 bp mRNA
DEFINITION Homo sapiens Immunoglobulin-like transcript 8 mRNA, complete cds.
ACCESSION AF041262
VERSION AF041262.1 GI:4104892
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
EXTRACTOR Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Mammalia: Primates: Catarrhini: Hominoidea: Homo.
REFERENCE 1 (bases 1 to 1395)
AUTHORS Colonna M.
TITLE Immunoglobulin-like transcript 8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1395)
AUTHORS Colonna M.
JOURNAL Immunoglobulin-like transcript 8
JOURNAL Submission
JOURNAL Submitted (07-JAN-1998) Basel Institute for Immunology, 487
JOURNAL, Grenzacherstrasse, Basel CH-4005, Switzerland
FEATURES
source
1..1395
/organism "Homo sapiens"
/db_xref "taxon:9606"
/chromosome "19"
/contig "513"
/cell "T89"
/clone "13"
1..1395
/function "activating receptor"
/notes "Immunoglobulin superfamily member: IL7R"
/codon_start 1
/product "Immunoglobulin-like transcript 8"
/old_locus "AF041262.1"
/old_xref "GI:4104892"
/translation "MPTLTATVTCIGLSLQPRTHVAGPFPKPTLMAEPGVSIVSGSP
VITICGSLSEAGVRYLKESGSEPMRNPLTPKNAKRSIPSTIHNGHYTCVYS
SAGVSESDALELVTAGVAFPTSLAPSVVAGSNNTLQCGSGKHPQFLAKESG
HQLPRTLSQGLHSGVQALPVPVPGVSHNRKTCYTYVWTPVPSWISDLELP
SGVSKRSILTLQGVLAQGLSLQCGSVQVQRYVLYKESKHPQFPQDQGL
SGVSKRSILTLQGVLAQGLSLQCGSVQVQRYVLYKESKHPQFPQDQGL
VASENNTLQCGSGKHPQFLAKESG
TRCYGSSPPLLSPEPRLLEWVSNAHQDYRNLKRNAGLIVYTLQILTR
AQSQRNPDQARR"
BASE COUNT 292 a 463 c 381 g 259 t
ORIGIN

```

```

alignment_scores:
  Quality: 94.50 Length: 126
  Ratio: 1.524 Gaps: 5
  Percent Similarity: 49.206 Percent Identity: 28.571
alignment_block:
  us-09-471-276-831 x AF041262

```

Align seq 1/1 to: AF041262 from: 1 to: 1395

```

1 MetSerMetLeuValValPheLeuLeuLeuTrkGlyValThrTrpGlyTr 17
  |||||
1 ATGACCCCATCTCCAGAGCTGATGATCTCTGCGAGCTGAGCTGGAGCC 50
  |||||
17 GvalTrkGlyAlaAlaLeuPheTrpGlyThrGln***SerLeuTrpAla 34
  |||||
51 CAGACCCACAGCTAGCCAGAGGCGCTCCACAGACCCGCTCGGGGCTG 100
  |||||
34 Lu.....SerGlnLys***LeuLysTrleu 42
  |||||
101 ACCACAGCTCTGTACACTGAGGAGGAGCCGTGACATCTGCTGAC 150
  |||||
43 GylGlnCysAspAla.....AspValProGlyProProGly 54
  |||||
151 CGAGAGCTGACAGCCAGAGTACAGCAGCATGATTAAGAGCAGACCC... 198
  |||||
54 YAspSerArGleuProAlaValGlnGluTrpGlyAlaGlnGluProVal 71
  |||||
199 .....GlnCysTrkGAGACAGATACAGCTG... 225
  |||||
71 IsLeuAspSerProAlaLeu1sleu1p1heuleu1thrc1yapTrp 87
  |||||
226 .....GACCCAGAGACAGCCGATTCCTCATTCCATCCATCCATTAACA 267
  |||||
88 Gln.....GlytGlytArGcYArGserGlyleuSerThrc1ytr 101
  |||||
268 GACGACACATGCGGAGAGATACCGCTGACATTCATACGCTGACAGCTG 317
  |||||
101 P***GlnleuSerLysleu1sleu1p1heuleu1thrc1yap 110
  |||||
318 CTCAGAGCCAGCAGCAGCCCTGCGACTG 345
  |||||
seq_name: gb_pf:AF322217
seq_documentation_block: 1745 bp mRNA
LOCUS AF322217
DEFINITION Rattus norvegicus Inhibin binding protein short isoform mRNA.
complete cds; alternatively spliced.
ACCESSION AF322217
VERSION AF322217.1 GI:18108899
KEYWORDS
Rattus norvegicus
ORGANISM Rattus norvegicus
EXTRACTOR Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae:
REFERENCE 1 (bases 1 to 1745)
AUTHORS Bernard D.J. and Woodruff, T.K.
TITLE Inhibin binding protein in rats: Alternative Transcripts and
Mol. Endocrinol. 15 (4): 654-667 (2001)
MEDLINE 21168090
PUBMED 11266515
REFERENCE 2 (bases 1 to 1745)
AUTHORS Bernard D.J. and Woodruff, T.K.
TITLE Direct Submission
JOURNAL Submitted (2000) Neurophysiology and physiology, Northwestern
University, 2153 North Campus Drive, Evanston, IL 60208, USA
JOURNAL Location/Qualifiers
source
1..1745
/organism "Rattus norvegicus"
/db_xref "taxon:10116"
170..871

```

```

/note="Inhib-S; alternatively spliced"
/codon_start=1
/product="inhibin binding protein short isoform"
/protein_id="AAK4084.1"
/db_xref="GI:11810500"
/translation="MALLCTHINRGSTNAVAVSGPFLIATSSNPQPMEN
ITLTKSSPSHYSSKLLKNDQNTQHTHPYKTSQSPFGLATSTSTLRICTYWR
KMSKSPKLELEAFQGLRPPIEMIOATETPPGCGVNVNLCHEMIDQIVLFFRECYT
EPIDYVPPGTAIVSLIDMLAENREVIYICRTHOVLPTLWSEPSNPLKLVAGGCAH
GCMILITVPIQIMAG"
BASE COUNT      481 a      436 c      345 g      483 t
ORIGIN
alignment_scores:
  Quality: 93.50      Length: 130
  Ratio: 1.438      Gaps: 8
  Percent Similarity: 50.000      Percent Identity: 34.615
alignment_block:
  US-09-471-276-831 x AF332217 ..
  Align seg 1/1 to: AF332217 from: 1 to: 1745
      3 MetLeuValAlaPheLeuLeuThrPdlGlyValThrTP.....GL 16
          ||||| .....||| ||||| |||
      173 ARGCTTCGGACCTTCACCTCTTGTGCTTCCTGCATTCGCTCATCGCG 222
          ||||| .....||| ||||| |||
      16 yProValThrGluAlaAlaIlePheTyGluThrGlu***SerLeuTPA 33
          ||||| .....||| ||||| |||
      223 TATGACTTCATGCGACGACGATG.....GACTGTCAACGACGACTTATGA 266
          ||||| .....||| ||||| |||
      33 IAGLSerGluHis**LeuLySThrLeuGlyGlnGlyAspAlaAspVal 49
          ||||| .....||| ||||| |||
      267 TACGATCCACACTACCC.....GCTCTCTCTCTCTCTCTCTCTCTCT 282
          ||||| .....||| ||||| |||
      50 ProGlyProGlyProGlyAspSerArg.....LeuProAla 60
          ||||| .....||| ||||| |||
      283 CCAGGCCCCCTGGAGACATCACTCTTGTCGAAAGAGCCCTCTCGGG 332
          ||||| .....||| ||||| |||
      60 aValGlnGluThrPdlGlyAlaGlnGluProValHis.....LeuAsp 74
          ||||| .....||| ||||| |||
      333 TATCCAGACGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
          ||||| .....||| ||||| |||
      74 er.ProAlaIleLys.....HisGluPheLeu.....LeuThr 84
          ||||| .....||| ||||| |||
      383 CATCTCTCTTACACAGACCTTCACAACTTCATTCCTCATAGTGCCTTAC 432
          ||||| .....||| ||||| |||
      84 rGlyAspThrGlnGluValArgTyArgCysArgSerGlyLeuSerThrGlyT 101
          ||||| .....||| ||||| |||
      433 TACGTGCTCAATACAGGCTGTTTATGATGATCTGCTGCTGCTGCTGCT 482
          ||||| .....||| ||||| |||
      101 PR***GlnLeuSerGlyLeuGlnGluLeuThrGlnGlyPro 113
          ||||| .....||| ||||| |||
      483 GGTGAGACCCAGATTAATCTCTAGAGCTAGAAAGACCA 520
seq_name: gb-pr:AF031554
seq_documentation_block:
  LOCUS      AF031554
  DEFINITION Homo sapiens clone 17.18 Immunoglobulin-like transcript 5 mRNA,
  complete cds.
  ACCESSION AF031554
  VERSION AF031554.1 GI:2655642
  KEYWORDS
  SOURCE
  ORGANISM
  Eukaryota; Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
  REFERENCE
  1 (bases 1 to 2066)
  Colonna,M., Navarro,F., Bellon,T., Llano,M., Garcia,P.,
  Samaridis,J.J., Angran,L., Cella,M. and Lopez-RotcL.M.
  TITLE
  A candidate inhibitory receptor for major histocompatibility complex

```

```

class I molecules on human lymphoid and myelomonocytic cells
J. Exp. Med. 186 (11), 1809-1818 (1997)
JOURNAL
MEDLINE
98044246
REFERENCE
2 (bases 1 to 2066)
Colonna,M.
AUTHORS
Mammalian B-cell lineage
JOURNAL
Submitted (127-OCT-1997) Basel Institute for Immunology, 487
Grenzacherstrasse, Basel 4005, Switzerland
FEATURES
SOURCE
1..2066
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/contig="AF031554.1"
/cell_line="myelomonocyte"
/clone="17.18"
1..1901
/note="Inhibitory receptor; ILT5"
/codon_start=3
/product="Immunoglobulin-like transcript 5"
/protein_id="AAC31591.1"
/translation="MPPALITALLCTGLSTLGPRTTQVAGAPPKPTLAEFGSVISNCS
PVTIMKCSLEADQYDOLKESGPEMDQRNNRLPRKNAKRSIPSTOHNGRTRCHYS
SGAKSEPSDPLLEWTGATSKLTLSLPSVVAASQNTLACSOGRYHNFLMKCEYS
HOLRPRLTSDQILSGTQVALPVGQPNVSHNRFTCYUYVNTPOVNSPDPLELP
SVNSKRPSLTLTGQPVLAHQSLTLGCGSDVGRVRLYKEDERQFLDRPQDQNGCL
SQANRPLTGVSPNSGQYKCGQANRCSKMSAPSDPLNTLMAQDLYTQVLSAQRPT
NSQKSGVSSSSQNDQNDSEPLTLEDAKESLQKQVQVQVQVQVQVQVQVQVQV
TTSQVSSSSQNDQNDSEPLTLEDAKESLQKQVQVQVQVQVQVQVQVQVQV
ATVLLPFLPFLLELRSHKRSKRSQKRSQKRSQKRSQKRSQKRSQKRSQKRSQK
QEBNLVAVAKDTQSEEDVLEDSQSFHUEDQPAVTVAAPVKHSSPRKMSPPSLSGE
FLDTRDQVYEDHQMDTAAESASQDVYTAQLISLILRRATEPPSPQDEBPAPS
IYATLAIH"
BASE COUNT      451 a      691 c      565 g      359 t
ORIGIN
alignment_scores:
  Quality: 93.50      Length: 126
  Ratio: 1.533      Gaps: 5
  Percent Similarity: 48.413      Percent Identity: 30.159
alignment_block:
  US-09-471-276-831 x AF031554 ..
  Align seg 1/1 to: AF031554 from: 1 to: 2066
      1 MetSerMetLeuValAlaPheLeuLeuThrPdlGlyValThrTPGlyPr 17
          ||||| .....||| ||||| |||
      3 ARGAGCCGCCGCTGACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 52
          ||||| .....||| ||||| |||
      17 oValThrGluAlaAlaIlePheTyGluThrGln***SerLeuTPAAla 34
          ||||| .....||| ||||| |||
      53 CAGACCCCGGTGCAGCGGCCCTCTCCCAAAACCCACCTCTGCGGCG 102
          ||||| .....||| ||||| |||
      34 Lu.....SerGluHis**LeuLySThrLeu 42
          ||||| .....||| ||||| |||
      103 AGCCAGGCTGTGATGACCTGGAGGAGACCGGACGATCTGATGTCAG 152
          ||||| .....||| ||||| |||
      43 GLYGlyAspAla.....AspAlaProGlyProGly 54
          ||||| .....||| ||||| |||
      153 CGAGACCTGAGGCCGAGGACATCCACATGGATTAAGAGGAGAACCCA.. 200
          ||||| .....||| ||||| |||
      54 yAspSerAlaPheProAlaValGlnGluThrPdlGlyValThrTPGlyPro 71
          ||||| .....||| ||||| |||
      201 .....GAGCCGCGGACGACAAATACACCACTG. 227
          ||||| .....||| ||||| |||
      71 ILeuAspSerProAlaIleLeuGlnGluThrPdlGlyValThrTPGlyPro 87
          ||||| .....||| ||||| |||
      228 .....GAAACCAACAAAGGCGAGATCTCCATCCATCCATCATGACA 269
          ||||| .....||| ||||| |||
      88 Gln.....GlyArgTyArgCysArgSerGlyLeuSerThrGlyTyr 101

```



```

88 Gln..... GtYArGtYfArGcYsArGserGtYleuSerrHrGtYTC 101
||||| ..... ||||| .....
3 AGCAGCCCTGACGAGGAGATACGCGTCCGACATTCATTCACGCTCCAGCGCTG 319
101 p***GlnleuSerrtYsleuGlnleu 110
||||| ..... ||||| .....
320 GTCAAGACCCGACGACCCCTGGAGCTG 347

seq_name: gb-pr:AF009639

seq_documentation_block: 2063 bp mRNA 05-DEC-1997
LOCUS AF031555 Homo sapiens clone 17.23 Immunoglobulin-like transcript 5 protein
DEFINITION mRNA, complete cds.
ACCESSION AF009639
VERSION AF009639.1 GI:2662437
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Placentalia; Cetartiodactyla; Hominidae; Homo.
AUTHORS Colonna,M., Navarro,F., Bellon,T., Llano,M., Garcia,P.,
Samaridis,J., Angman,L., Celis,M. and Lopez-Bolet,M.
TITLE A common inhibitory receptor for major histocompatibility complex
class I molecules on human lymphoid and myelomonocytic cells
JOURNAL Med. 186 (11), 1809-1818 (1997)
MEDLINE 98044246
REFERENCE
AUTHORS Colonna,M.
TITLE 2 (bases 1 to 2063)
Direct Submission
Submitted (20-JUN-1997) Basel Institute for Immunology, 487
JOURNAL Grenzacherstrasse, Basel CH-4005, Switzerland
FEATURES
1..2063
location/Qualifiers
1..2063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/cell_type="myelomonocytic"
/clone="31"
/map="9q13.4"
J. Exp. Med. 186 (11), 1809-1818 (1997)
/note="Inhibitory receptor."
/codon_start=1
/product="Immunoglobulin-like transcript 5 protein"
/protein_id="AAC51894.1"
/translation="MTATLALICGLISGLPTFMKAGPPKFLTMAEGSVISKSGS
SAGSBSRPSLELWGTGFYFNKPTSLSDPVVASGNTLACSGKHIVHLMKEDP
HQLPRLDSGLSGFOALPPGVPFSIRHARPTCYTYTNTPMWHSPPDLLELP
SGVSRKLSLLTLCQPLVLAQSLITLQSGDVGVDRFLVYKEGRDFLQHPQDPAGL
SGAUNTIGPVSRYGQCYGAMLSSEMSAPSDLLILITGQIVDTYVLSAQPGPT
VASGEMNLQCSRGYFDFLRLKEGAHPLRLKSNHGAKTQAEPMFMSVSNAG
TRCTYSKSNPSPLSLSPEDLELWHSNGSSGLVYVPGPSTPELGRLELVLSVY
QRENLVYAKMTQSDVSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD
LPTDQGVREDRQMTAEASASDQVTVTAQNLTLRLKATPEPPSGCGPAPRSTI
VATLALH"
BASE COUNT 449 a 690 c 561 g 363 t
ORIGIN
Alignment.L.sources:
Quality: 92.50 Length: 126
Ratio: 1.542 Gaps: 5
Percent Similarity: 47.619 Percent Identity: 30.159
Alignment_block:
US-09-471-276-831 x AF009639 ..
Align beg 1/1 to: AF009639 from: 1 to: 2063

1 MelSerHelleuValValPheleuLeuLeuTrpGlyValThrTrpGly 17
||||| ..... ||||| .....
3 ATGACGCCCTGACGAGGAGATACGCGTCCGACATTCATTCACGCTCCAGCGCTG 52
17 ovalTheuLualValIleHethrYrGtYthGtYp***SerGlnTrpAlaG 34
||||| ..... ||||| .....
53 CAGCAGCCGATGATGACGACGACCCCTTCCCAACACCCGCTCTGGGCTG 102
34 Lu..... SerGlnIleH***ValYsThrLeu 42
||||| .....
103 ACCGACGCTGTGTGATACGCTG336GAGCCCTGACGCTGTGCTGAC 152
43 GluGlnCysAspAla..... AspValTrpGlyTrpProGly 54
||||| ..... ||||| .....
153 GCGACGCTGGAGGCCGACAGCTACCAATTCGATTAAGACGACACCCA... 200
54 YAspSerfArgLeuProAlaValGlnGlnuTrpGlyValGlnuTrpValH 71
||||| ..... ||||| .....
201 ..... GAGCCCTGGAGGAGATACGACATTCACGCTCCAGCGCTG 227
71 IseuAspSerProAlaIleValGlnGlnuTrpGlyLeuTrpGlyAspArg 87
||||| ..... ||||| .....
228 ..... GAACCCAGAAAGAACGCCGACATTCGACATTCACGCTCCAGCGCTG 269
88 Gln..... GtYArGtYfArGcYsArGserGtYleuSerrHrGtYTC 101
||||| ..... ||||| .....
270 CAGCAGCCCTGACGAGGAGATACGCGTCCGACATTCATTCACGCTCCAGCGCTG 319
101 p***GlnleuSerrtYsleuGlnleu 110
||||| ..... ||||| .....
320 GTCAAGACCCGACGACCCCTGGAGCTG 347

seq_name: gb-pr:AF031555

seq_documentation_block: 2063 bp mRNA 09-DEC-1997
LOCUS AF031555 Homo sapiens clone 17.23 Immunoglobulin-like transcript 5 mRNA,
DEFINITION complete cds.
ACCESSION AF031555
VERSION AF031555.1 GI:2665644
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Placentalia; Cetartiodactyla; Hominidae; Homo.
AUTHORS Colonna,M., Navarro,F., Bellon,T., Llano,M., Garcia,P.,
Samaridis,J., Angman,L., Celis,M. and Lopez-Bolet,M.
TITLE A common inhibitory receptor for major histocompatibility complex
class I molecules on human lymphoid and myelomonocytic cells
JOURNAL J. Exp. Med. 186 (11), 1809-1818 (1997)
MEDLINE 98044246
REFERENCE
AUTHORS Colonna,M.
TITLE 2 (bases 1 to 2063)
Direct Submission
Submitted (12-JAN-1997) Basel Institute for Immunology, 487
JOURNAL Grenzacherstrasse, Basel CH-4005, Switzerland
FEATURES
1..2063
location/Qualifiers
1..2063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.4"
/chromosome="19"
/cell_type="myelomonocytic"
/clone="31"
/map="9q13.4"
J. Exp. Med. 186 (11), 1809-1818 (1997)
/note="Inhibitory receptor; IL75"
/codon_start=3
/product="Immunoglobulin-like transcript 5*"
/protein_id="AAC51902.1"
/translation="MTATLALICGLISGLPTFMKAGPPKFLTMAEGSVISKSGS
PTIMCGSLDNEOTLWFKATLSSSPEDWKNRPFLPRKATNFSIPSNQHNHGHCTHTS

```

SAGSESPDPLELVTMGVFNKPTLSALPSPVVAAGGNTLRGCSQKGYHIFVLMKEE
HOLPRTLDLSQGLHSGFOALPVGCVGTPSHMRFTCTVYTATPWGSHPSDLEILP
SOVSRKPSGLTJTGPGVYLAFOGSLTJGCSGSDGVDRVYLAKEEDFLORPOOPAGL
SOANFTLTPSVNSGSDVRCYGAHNLSSSKAPBDLITLIGQITDVSLSAOGPT
SOSEBTELSQKGLHSGFOALPVGCVGTPSHMRFTCTVYTATPWGSHPSDLEILP
TYRCSKSSNBRPLTSPSPPLTLYOVCHSGSDPADPGPSTGALCEYVYLSQK
AFVILLFLLELLLRORSHKHTSDORTDIPORPAGATSPEDRGLIRRSSPAAY
OENLYAAVKDQSDVELDSQSDPHEDPOAVYVAPKISSPRRMASSPSSLGCF
LDTRKQROVEDKQMDTEAASASQDVTYAOJLSLTLRRKATEPPSOCEPAPPSI
VATLAIH"

BASE COUNT 447 a 691 c 563 g 362 t
ORIGIN

Alignment_scores:
Quality: 92.50 Length: 126
Ratio: 1.542 Gaps: 5
Percent Similarity: 47.619 Percent Identity: 30.159

alignment_block:
US-09-471-276-831 x AF031555 ..

Align seg 1/1 to: AF031555 from: 1 to: 2063

```

1 MeSerMetLeuValValPheLeuLeuLeuTrpGlyValThrProGlyPyr 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 ATGACCCGCCGCCCTCCACACGCCCTCTCCCTGCGCTAGCTGGAGGCC 52
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 ovaThrlGlnAlaIlePheTrpGluThrGln***SerLeuTrpAlaG 34
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 CAGGACCCGATGACGACGAGGCCCTCCGCCAAGCCACCTCGGGCTG 102
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 Lu.....SerGlnHis**LeuIysThrLeu 42
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 AGCCAGCCTCTGTGATCAGCTGGGGAGGCCCTGATCAGCATCTGGTACG 152
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 GlycylGlnAspAla.....AspValProGlyProProGly 54
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 GCGAGCCTGGAGGCCGACGATACCACTGATTAAGGAGGAGGCCCA 200
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 yAspSerArgLeuProAlaValGlnGluTrpGlyValGlnIleuProValH 71
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 .....GAGCCCTGGGACAGAAATAAACCCACG 227
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 IsoleuAspSerProAlaIleIysHisGlnPheLeuLeuThrGlyAspThr 87
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 .....GACCCACAGAACAGCCGACATTCCTCCATCCATCCATCAGCA 269
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 Gln.....GlyArgTyrArgCysArgSerGlyLeuSerThrGlyPyr 101
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 CAGGACCCGATGACGACGAGGATACCGCTCTCCCTGCGCTAGCTGGAGGCC 319
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 P***GlnLeuSerTrpGlyLeuLeuIleuLeu 110
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 GTCAGAGCCGATGACGACGAGGCCCTCCGAGACTG 347
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: gb.Pr.AF009635

seq_documentation_block: 2066 bp mRNA 05-DIC-1997
103 AGCCAGCCTCTGTGATCAGCTGGAGGCCCTGATCAGCATCTGGTACG 152
DEFINITION Homo sapiens clone 17.7 immunoglobulin-like transcript 5 protein
mRNA, complete cds.
ACCESSION AF009635
VERSION AF009635.1 GI:2662429
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
TAXONOMY Eukaryota; Chordata; Cetiartia; Vertebrata; Eulcolecotmi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Colonna M., Navarro F., Bellon T., Llano M., Garcia P.,
Samaridis J., Amman L., Cella M. and Lopez-Rotet M.
TITLE A common inhibitory receptor for major histocompatibility complex

class I molecules on human lymphoid and myelomonocytic cells
J. Exp. Med. 186 (11), 1809-1818 (1997)
MEDLINE 98044246
REFERENCE 2 (bases 1 to 2066)
NOTIONS Coloma H. Immun
JOURNAL Submission
Submitted (20-JUN-1997) Basel Institute for Immunology, 487
Grenzacherstrasse, Basel CH-4005, Switzerland
FEATURES
source
1..2066
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cel_ltype="mrg1omonocytic"
/clone="19"
/map="19q13.4"
3..1901
/note="Immunoglobulin superfamily; inhibitory receptor"
/codon_start=1
/product="immunoglobulin-like transcript 5 protein"
/protein_id="ACG2430.1"
/translation="MPAPATLALICIGSLGAPTRPMQAGPPPTLAAKPSVISMSP
VTICGSLAEQRIQDKEGSPENDNNLEPKNAKAFSTPSMTHAGRYKCHYS
SAGSESPDPLELVTMGVFNKPTLSALPSPVVAAGGNTLRGCSQKGYHIFVLMKEE
HOLPRTLDLSQGLHSGFOALPVGCVGTPSHMRFTCTVYTATPWGSHPSDLEILP
SOVSRKPSGLTJTGPGVYLAFOGSLTJGCSGSDGVDRVYLAKEEDFLORPOOPAGL
SOANFTLTPSVNSGSDVRCYGAHNLSSSKAPBDLITLIGQITDVSLSAOGPT
SOSEBTELSQKGLHSGFOALPVGCVGTPSHMRFTCTVYTATPWGSHPSDLEILP
TYRCSKSSNBRPLTSPSPPLTLYOVCHSGSDPADPGPSTGALCEYVYLSQK
AFVILLFLLELLLRORSHKHTSDORTDIPORPAGATSPEDRGLIRRSSPAAY
OENLYAAVKDQSDVELDSQSDPHEDPOAVYVAPKISSPRRMASSPSSLGCF
LDTRKQROVEDKQMDTEAASASQDVTYAOJLSLTLRRKATEPPSOCEPAPPSI
IYATLAIH"

CDS

BASE COUNT 447 a 693 c 567 g 359 t
ORIGIN

alignment_scores:
Quality: 92.50 Length: 126
Ratio: 1.542 Gaps: 5
Percent Similarity: 47.619 Percent Identity: 30.159

alignment_block:
US-09-471-276-831 x AF009635 ..

Align seg 1/1 to: AF009635 from: 1 to: 2066

```

1 MeSerMetLeuValValPheLeuLeuLeuTrpGlyValThrProGlyPyr 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 ATGACCCGCCGCCCTCCACACGCCCTCTCCCTGCGCTAGCTGGAGGCC 52
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 ovaThrlGlnAlaIlePheTrpGluThrGln***SerLeuTrpAlaG 34
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 CAGGACCCGATGACGACGAGGCCCTCCGCCAAGCCACCTCGGGCTG 102
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 Lu.....SerGlnHis**LeuIysThrLeu 42
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 AGCCAGCCTCTGTGATCAGCTGGAGGCCCTGATCAGCATCTGGTACG 152
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 GlycylGlnAspAla.....AspValProGlyProProGly 54
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 GCGAGCCTGGAGGCCGACGATACCACTGATTAAGGAGGAGGCCCA 200
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 yAspSerArgLeuProAlaValGlnGluTrpGlyValGlnIleuProValH 71
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 .....GAGCCCTGGGACAGAAATAAACCCACG 227
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 IsoleuAspSerProAlaIleIysHisGlnPheLeuLeuThrGlyAspThr 87
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 .....GACCCACAGAACAGCCGACATTCCTCCATCCATCCATCAGCA 269
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 Gln.....GlyArgTyrArgCysArgSerGlyLeuSerThrGlyPyr 101
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


[illegible]

```

misc_feature      comp|ement(2022..2111)
                  /note="SRAL1.2 excellent exon, frame 2"
misc_feature      complement(join(2084..2173,3273..3512))
                  /note="92% identity mouse EST m33404.r1"
repeat_region     2084..xref|dbEST:AA169004*
misc_feature      comp|ement(3317..3460)
                  /note="SRAL1.2 excellent exon, frame 2"
BASE COUNT      915 a 1297 c 1466 g 1077 t 1 others
ORIGIN
alignment_scores:
  quality: 92.50      length: 107
  ratio: 1.652        gaps: 4
  percent identity: 52.336   percent identity: 27.103

alignment_block:
  US-09-471-276-831 x AC005590/rev ..
  Align seg 1/1 to reverse of: AC005590 from: 1 to: 4756

10 LeuTPRGLY.....ValThrPGLYProValThrGLYAlaAla11 23
   :::::|||||
1420 GGTGGGGGTGGAGAGATGCTCCCGGGCCCTCCAGCCCTCCCTCT 1371
   :::::|||||
23 ePheryrGLuThrGLi***SerLeuTrPAlaGLuSerGLuHLS***LeuL 40
   :::::|||||
1370 TCCTTCCCAACAGCTCTTCACTCCCTCCCTCCCTCTGGGAACCTC 1322
   :::::|||||
40 ySthrLeuGLyGLiCysAspAlaAspValProGLYProProGLYAsp 56
   :::::|||||
1321 .....TGCATGCTGCTCCGACCCCTCCGCTCCGCTCCATGCA 1286
   :::::|||||
57 ArgLeuProAlaValGLuThrPGLYAlaGLuThrPGLYAlaGLu 73
   :::::|||||
1285 CATTCCCAAGAGTCAACCCAGTGTGGCTCC.....ATTCACTGGCA 1245
   :::::|||||
73 pSerProAla11LeuYsHLSInpHe.....L 82
   :::::|||||
1244 GGAGCTGCTGCTCCAGTGGAGCCCTCTCCACCTCCCTCTCTTTCGAC 1195
   :::::|||||
82 euLeuThrGLYAspThrGLYAlaGLuThrPGLYAspGLYAspGLYLeu 98
   :::::|||||
1194 TCCTTCCGGAACAGTGGAGGCACTTCTCTCTCCAGGAGCAACGCA 1145
   :::::|||||
99 ThrGLYTrP**GlnLeuSer 105
   :::::|||||
1144 GTATTATGACACACTTAC 1124
seq_name: gb_hlg.AC0041031

seq_documentation_block:
LOCUS      AC0041031      68207 bp      DNA      HTG      11-APR-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-684C2 map 2, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION  AC0041031.1 GI:7534234
VERSION    AC0041031.1
KEYWORDS   HTG; RHOS; PHASISO.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eulambda; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 68207)
            Birren,B., Linton,L., Nusbaum,C. and Lander,E.
            The human genome: initial sequencing and assembly.
            1994. Science 266: 1355-1371.
            2 (bases 1 to 68207)
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,M., Allen,N.,
            Boguski,M., Bult,C., Burt,A., Brown,A., Burkett,C.,
            Campione,A., Castle,A., Chao,P., Collins,S.,

```

```

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 120 Charles Street, Cambridge, MA 02141, USA
All requests were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Center project name: 1962
Center clone name: 684_C_2
-----
* NOTE: This record contains 77 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. The reads are not necessarily in order
* of overlapping clones that may be gene-rich and allows
* identifying relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
760 855: contig of 769 bp in length
860 1626: contig of 767 bp in length
1627 1726: gap of 100 bp
1727 2507: contig of 781 bp in length
2508 2607: gap of 100 bp
2608 3396: contig of 789 bp in length
3397 3496: gap of 100 bp
3497 4279: contig of 783 bp in length
4280 4379: gap of 100 bp
4380 5137: contig of 758 bp in length
5138 5237: gap of 100 bp
5238 5988: contig of 761 bp in length
5989 6098: gap of 100 bp
6099 6873: contig of 781 bp in length
6880 6979: gap of 100 bp
6980 7765: contig of 783 bp in length
7766 7865: gap of 100 bp
7866 8633: contig of 768 bp in length
8634 8733: gap of 100 bp
8734 9515: contig of 782 bp in length
9516 9615: gap of 100 bp
9616 10399: contig of 784 bp in length
10400 10493: gap of 100 bp
10494 11289: contig of 790 bp in length
11290 12179: contig of 790 bp in length
12180 12279: gap of 100 bp
12280 13046: contig of 767 bp in length
13047 13146: gap of 100 bp

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2002, 16:46:53 : Search time 31.95 seconds

(without alignments) 576,849 Million cell updates/sec

File: us-09-471-276-831

Protein score: 1 NENLWVFLIMGVTVMGVYVE.....ILEITGRVYLACSLALDGAS 126

Sequence: Gapop 10.0, Gapext 0.5

Scoring table: BIOSUM62

473505 seps, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	130.5	19.7	513	11	Q9EPH1 rat mus norv
2	90.5	13.7	631	4	Q95P11
3	87.5	13.2	631	4	Q15471 mus sapien
4	84.5	12.8	448	4	Q15468 mus sapien
5	84.5	12.8	448	4	Q15468 mus sapien
6	84.5	12.8	448	4	Q15468 mus sapien
7	78.5	11.5	652	4	Q75025 mus sapien
8	78	11.8	192	12	Q79089 human immun
9	76	11.5	1327	4	Q92HP2
10	75.5	11.4	244	2	Q92HP2
11	75.5	11.4	635	11	Q55002
12	75	11.3	192	12	Q90719
13	75	11.3	428	4	Q9AKK4
14	75	11.3	428	4	Q9AKK4
15	75	11.3	647	4	Q9AKK4
16	74.5	11.3	648	4	Q9P209
17	74.5	11.3	663	11	Q70434
18	74.5	11.3	680	11	Q55001
19	74	11.2	192	12	Q9WPD4

20	74	11.2	289	4	Q75018	Q75018 homo sapien
21	74	11.2	650	4	Q75024	Q75024 homo sapien
22	72.5	11.0	187	10	Q9LYK0	Q9LYK0 arabidops
23	72.5	11.0	192	12	Q91076	Q91076 human immun
24	72	10.9	192	12	Q92HP2	Q92HP2 mus sapien
25	72	10.9	104	12	Q73438	Q73438 human immun
26	72	10.9	192	12	Q73410	Q73410 human immun
27	72	10.9	192	12	Q73413	Q73413 human immun
28	72	10.9	192	12	Q73439	Q73439 human immun
29	72	10.9	192	12	Q910D9	Q910D9 human immun
30	72	10.9	475	5	Q9VW92	Q9VW92 drosophila
31	71.5	10.8	485	2	Q9W034	Q9W034 pseudomonas
32	71	10.7	192	12	Q73435	Q73435 human immun
33	71	10.7	192	12	Q73435	Q73435 human immun
34	71	10.7	192	12	Q92HP2	Q92HP2 mus sapien
35	71	10.7	288	7	Q45087	Q45087 bos taurus
36	71	10.7	643	5	Q45087	Q45087 caenorhabd
37	70	10.6	147	12	Q73436	Q73436 human immun
38	70	10.6	192	12	Q73434	Q73434 human immun
39	70	10.6	192	12	Q92HP2	Q92HP2 mus sapien
40	70	10.6	192	12	Q92HP2	Q92HP2 mus sapien
41	69.5	10.5	728	12	Q9Y0X0	Q9Y0X0 rapid hepe
42	69.5	10.5	841	11	Q54959	Q54959 mus muscu
43	69	10.4	192	12	Q71262	Q71262 human immun
44	69	10.4	192	12	Q71966	Q71966 human immun
45	69	10.4	192	12	Q71971	Q71971 human immun
46	69	10.4	192	12	Q73411	Q73411 human immun
47	69	10.4	192	12	Q73411	Q73411 human immun
48	69	10.4	192	12	Q73411	Q73411 human immun
49	69	10.4	192	12	Q79115	Q79115 human immun
50	69	10.4	192	12	Q79128	Q79128 human immun
51	69	10.4	192	12	Q80237	Q80237 human immun
52	69	10.4	192	12	Q81977	Q81977 human immun
53	69	10.4	254	12	Q88704	Q88704 porcine rep
54	69	10.4	324	5	Q9T002	Q9T002 hydra post
55	69	10.4	488	1	Q75018	Q75018 mus muscu
56	69	10.4	488	1	Q75018	Q75018 mus muscu
57	68.5	10.4	244	10	Q9P0D2	Q9P0D2 arabidops
58	68.5	10.4	831	11	Q9ER82	Q9ER82 mus muscu
59	68.5	10.4	831	11	Q9ER82	Q9ER82 mus muscu
60	68	10.3	185	12	Q73433	Q73433 human immun
61	68	10.3	192	12	Q73431	Q73431 human immun
62	68	10.3	192	12	Q73432	Q73432 human immun
63	68	10.3	192	12	Q73432	Q73432 human immun
64	68	10.3	192	12	Q73432	Q73432 human immun
65	68	10.3	192	12	Q73432	Q73432 human immun
66	68	10.3	192	12	Q73432	Q73432 human immun
67	68	10.3	192	12	Q73432	Q73432 human immun
68	68	10.3	192	12	Q73432	Q73432 human immun
69	68	10.3	192	12	Q73432	Q73432 human immun
70	68	10.3	192	12	Q73432	Q73432 human immun
71	68	10.3	192	12	Q73432	Q73432 human immun
72	68	10.3	192	12	Q73432	Q73432 human immun
73	68	10.3	192	12	Q73432	Q73432 human immun
74	68	10.3	192	12	Q73432	Q73432 human immun
75	68	10.3	192	12	Q73432	Q73432 human immun
76	68	10.3	192	12	Q73432	Q73432 human immun
77	68	10.3	192	12	Q73432	Q73432 human immun
78	68	10.3	192	12	Q73432	Q73432 human immun
79	68	10.3	192	12	Q73432	Q73432 human immun
80	68	10.3	192	12	Q73432	Q73432 human immun
81	68	10.3	192	12	Q73432	Q73432 human immun
82	68	10.3	192	12	Q73432	Q73432 human immun
83	68	10.3	192	12	Q73432	Q73432 human immun
84	68	10.3	192	12	Q73432	Q73432 human immun
85	68	10.3	192	12	Q73432	Q73432 human immun
86	68	10.3	192	12	Q73432	Q73432 human immun
87	68	10.3	192	12	Q73432	Q73432 human immun
88	68	10.3	192	12	Q73432	Q73432 human immun
89	68	10.3	192	12	Q73432	Q73432 human immun
90	68	10.3	192	12	Q73432	Q73432 human immun
91	68	10.3	192	12	Q73432	Q73432 human immun
92	68	10.3	192	12	Q73432	Q73432 human immun

CC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
OC	NCBI_TaxID:9606;	
RN	111	
RP	SEQUENCE FROM N.A.	
DR	Ensembl: ENSEMBL:ENSG00000187633.5	
CC	1 - SIMILARITY: 0.0-0.0(1997)	
CC	1 - SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN	
DR	EMBL: AF025533; AAB87667.1; *	
DR	HSSP: P43626; 1NR.	
DR	InterPro: IPR003599; Ig	
DR	InterPro: IPR003600; Ig_Like	
DR	InterPro: IPR003601; Ig_Like	
DR	InterPro: IPR007234; Lipase	
DR	Pfam: PF00047; 1g; 3	
DR	SMART: SM00409; 1G; 1.	
DR	SMART: SM00410; 1G_Like; 2.	
DR	PROSITE: PS00120; LIPASE_SER; UNKNOWN.1.	
SO	SEQUENCE 631 AA; 69226 MW; E77E1EDB042F9C45 CIRC64;	
Query Match	13.7%; Score 90.5; db 4; Length 631;	
	Best Local Similarity 28.3%; Pval No. 0.096;	
	Matches 39; Conservative 10; Mismatches 38; Indels 51; Gaps	
Oy	1 KSMVFLIMLVGVWVPTGTAIFETQSLAAESEHXLKTLQCDVDVGRHGSRYRA 60	
Db	1 PRLALALALCLLDLPDPRTFNQAGPRFKPTLME-----PQ-----S 32	
Oy	61 VOEMG-----AOE-----PYHLD-----SPAINQFLLTGDDP--GRY 97	
Db	38 VISMGSPTTVMQGSLSAEQVQLKDEGSPEDLNHNPFLPKNKAFTSPKOHNAGRYA 97	
Oy	93 CRGSLSTGMAKSLDEL 110	
Db	98 CHYSGMNSRSTPDEL 115	
RESULT 3		
ID	015471	PRELIMINARY; PRT; 631 AA.
AC	015471	
DT	01-JAN-1998 (TRIMBREL 05, Created)	
DT	01-JAN-1998 (TRIMBREL 05, Last annotation update)	
DT	01-JUN-2001 (TRIMBREL 17, Last annotation update)	
DE	MONOCYTE INHIBITORY RECEPTOR PRECURSOR.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
NC	NCBI_TaxID:9606;	
RP	SEQUENCE FROM N.A.	
DR	Ensembl: ENSEMBL:ENSG00000187633.5	
CC	1 - SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN	
DR	EMBL: U01928; AAB6668.1; *	
DR	HSSP: P43626; 1NR.	
DR	InterPro: IPR003599; Ig	
DR	InterPro: IPR003600; Ig_Like	
DR	InterPro: IPR003601; Ig_Like	
DR	InterPro: IPR007234; Lipase	
DR	Pfam: PF00047; 1g; 3	
DR	SMART: SM00409; 1G; 2.	
DR	SMART: SM00410; 1G_Like; 1.	
DR	PROSITE: PS00120; LIPASE_SER; UNKNOWN.1.	
FT	CHAIN 1 23	POTENTIAL INHIBITORY RECEPTOR.
FT	SIGNAL 1 631	POTENTIAL INHIBITORY RECEPTOR.
SO	SEQUENCE 631 AA; 69385 MW; 61CF1967E08D79FA CIRC64;	
Query Match	13.7%; Score 87.5; DB 4; Length 631;	

Best Local Similarity 27.5%; Pred. No. 0.21;
Matches 38; Conservative 11; Mismatches 38; Indels 51; Gaps 6;

OY 1 MSNLVFLILMGVMPVTEALIFETQXSLAASESHIKATLGGCDADVPAPRGDSRPA 60
DB 1 MRLPALCLGLSLGPRRVOAGPPPPYLAHME-----TG-----S 37
OY 61 VOEGK-----ADE-----PHLID-----SPAIRHGFLLTGDQ---GARY 92
DB 38 VISMGSPTVIMCGSLSLMDGSHYLRDKESPSPLDRNPLFNKARFSISMTENHAGYR 97
OY 93 CRSGSLSTGNQSLSKLLR 110
DB 98 CHYSSAKGSEPSDPLR 115

RESULT 4
015468 PRELIMINARY: PRT; 448 AA.

ID 015468
AC 015468: 1998 (TREMblrel). 05. Created
DT 01-JAN-1998 (TREMblrel). 05. Last sequence update
DE 01-JUN-2001 (TREMblrel). 17. Last annotation update
DE MONOCYTE INHIBITORY RECEPTOR PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID=9606;
RA NCBI_TaxID=9606;
RM SEQUENCE FROM N.A.
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.

DR EMBL: 091925: AAB68665.1. -
DR HSSP: P43626: INKR. 19. Like.
DR InterPro: IPR003006: 19_MHC.
DR Pfam: PF00047: 19; 2. 19_MHC.
DR SMART: SM00410: IG_Like; 2.
KM Signal. 23
FT SIGNAL 24
FT CHAIN 1 448
FT SEQUENCE 448 AA; 49297 MW; 76D1E24A92EA1399 CRC64.

Query Match 12.8%; Score 84.5; DB 4; Length 448;
Best Local Similarity 27.9%; Pred. No. 0.31;
Matches 34; Conservative 11; Mismatches 44; Indels 33; Gaps 5;

OY Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID=9606;
RA NCBI_TaxID=9606;
RM SEQUENCE FROM N.A.
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.

DR EMBL: AF025534: AAB87668.1. -
DR HSSP: P43626: INKR.
DR InterPro: IPR003599: 19.
DR InterPro: IPR003600: 19_Like.
DR Pfam: PF00047: 19; 2. 19_MHC.
DR SMART: SM00410: IG_Like; 2.
DR SMART: SM00410: IG_Like; 2.
SO SEQUENCE 448 AA; 49325 MW; 76D1E0B7AD3A1399 CRC64;

Query Match 12.8%; Score 81.5; DB 4; Length 590;
Best Local Similarity 27.9%; Pred. No. 0.89; 55; Indels 27; Gaps 6;

OY 1 MSNLVFLILMGVMPVTEALIFETQXSLAASESHIKATLGGCDADVPAPRGDSRPA 60
DB 1 MRLPALCLGLSLGPRRVOAGPPPPYLAHME-----TG-----S 37

RESULT 6
075023 PRELIMINARY: PRT; 590 AA.

ID 075023
AC 075023:
DT 01-NOV-1998 (TREMblrel). 08. Created
DT 01-NOV-1998 (TREMblrel). 08. Last sequence update
DE 01-JUN-2001 (TREMblrel). 17. Last annotation update
DE LEUCOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID=9606;
RA NCBI_TaxID=9606;
RM SEQUENCE FROM N.A.
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.

DR EMBL: AF025534: AAB87668.1. -
DR HSSP: P43626: INKR.
DR InterPro: IPR003599: 19.
DR InterPro: IPR003600: 19_Like.
DR Pfam: PF00047: 19; 2. 19_MHC.
DR SMART: SM00410: IG_Like; 2.
DR SMART: SM00410: IG_Like; 2.
SO SEQUENCE 590 AA; 64125 MW; 8E7AB249F12E44CD CRC64;

```

OY 48 DYPGPGDSRLPAVQENQAEPRVILDSFALIKIOPFLITGTCGRYKNSGLSTGKQSLK 107
DB 61 DKCGLPMAKRNQLLEPGANAFHITSTVY-----DSACRYRCYETEPACSESPDP 112
OY 108 LEL--TG---PXYLA 117
DB 113 LELVATGFAVPEPLLA 128

RESULT 7
ID 075025 PRELIMINARY: PRT: 652 AA.
AC 075025:
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE LECOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-1.
OS Homo sapiens (human).
OC Muscularia, Euteleostomi, Chordata, Vertebrata, Euteleostomi:
OC Mammalia, Euteleostomi, Primates, Catarrhini, Hominoidea, Homo.
NCBI_TaxId:9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Cosman D., Fanger N., Borges L., Kubla M., Chin M., Peterson L.,
RA Hsu M.L. (JUL-1997) to the EMBL/Genbank/CDDB databases.
NCBI_Study: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
EMBL: AF009221; AAB63522.1; -.
DR HSRP: P43626; INKR.
DR InterPro: IPR003599; 19.
DR InterPro: IPR003600; 19_Like.
DR RefSeq: NM002413.3.
DR SMART: SM00409; 16; 2.
DR SMART: SM00410; 16_Like; 1.
DR SMART: SM00410; 16_Like; 1.
SQ SEQUENCE 652 AA: 71018 MW: 99AF0P21A2945D CXC64;

Query Match 11.9%; Score 78.5; DB 4; Length 652;
Best Local Similarity 24.8%; Pred. No. 2.1; 57; Indels 53; Gaps 7;
Matches 41; Conservative 16; Mismatches 16;

OY 1 MSMLVYLLMGVNGYRPAIFYEOKSIAMSESHXIKTLQGDNDVPGPDSRLPA 60
DB 1 MPEITVILICGLSLGPRTHVAGHRIKPLMAEGSVI-----TQGSPTVTKQCG 51
OY 61 VQE-----WGAEPRVILDSFALIKIOPFLITGTCGRYKNSGLSTGKQSLK 102
DB 52 GQGTQVYALYREKKTALYIKTRIGEL--VKQGPRTSTWEMGRKRCITGSDTACRS 108
OY 103 QLSKLELT-----CRVLAC--SLADGAS 126
DB 109 ESSSDLELVYGVATIKPLTSAQPSFVNSGVNLIQDQSOVAFDPSFS 155

RESULT 8
ID 079089 PRELIMINARY: PRT: 192 AA.
AC 079089:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VIP (VIRAL INFECTIVITY FACTOR).
OS Homo sapiens (human).
OC Muscularia, Euteleostomi, Chordata, Vertebrata, Euteleostomi:
OC Mammalia, Euteleostomi, Primates, Catarrhini, Hominoidea, Homo.
NCBI_TaxId:9606;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN-PATIENT A30.
RA MEDLINE:94303241; PubMed:8030283;
RA Wieland U., Hartmann H., Salzhager B., Eggers H.J.,

```

```

RA Kuehn J.E.;
RT "In vivo genetic variability of the HIV-1 vif gene.";
RL Virology 2003;43-51(1994).
DB 23 LMDQPELILNESN-----DQANWEN--ITLIL-CASRSIRSKTLLKLP 113
OY 23 IFEYTONSIAMSESHXIKTLQGDNDVPGPDSRLPAVQENQAEPRVILDSFALIKI 107
DB 23 LMDQPELILNESN-----DQANWEN--ITLIL-CASRSIRSKTLLKLP 113
OY 75 -----PAIKH---OFI---LTGDPGQRYGRSGSTGQWQMSLILATPC 113
DB 66 TQMTWIPRSHRTQVSPLIGALTSTHAGLVRCQWKEBTQMSKPSVLELEAP 117

RESULT 9
ID 015070 PRELIMINARY: PRT: 1327 AA.
AC 015070:
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE KIA00364.
OS Homo sapiens (human).
OC Muscularia, Euteleostomi, Chordata, Vertebrata, Euteleostomi:
OC Mammalia, Euteleostomi, Primates, Catarrhini, Hominoidea, Homo.
NCBI_TaxId:9606;
RN (1)
RP SEQUENCE FROM N.A.
RA TISSUE-BRAIN.
RX MEDLINE:97349984; PubMed:9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohtsu M., Seki N., Miyajima N.,
RA Prediction of the coding regions of newly identified human genes. V.
RT The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro.
CC 1- SIMILARITY TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
EMBL: AF002262; BAA20819.1; -.
DR HSRP: P43626; INKR.
DR InterPro: IPR003598; 19_C2.
DR InterPro: IPR003600; 19_Like.
DR InterPro: IPR003600; 19_Like.
DR Pfam: PF00047; 19; 12.
DR SMART: SM00408; 16; 2.
DR SMART: SM00410; 16_Like; 1.
DR SMART: SM00410; 16_Like; 1.
SQ SEQUENCE 1327 AA: 147971 MW: 736C699FEC94D2E1 CXC64;

Query Match 11.5%; Score 76; DB 4; Length 1327;
Best Local Similarity 26.8%; Pred. No. 8.8; 34; Indels 38; Gaps 6;
Matches 30; Conservative 10; Mismatches 34;

OY 23 IFEYTONSIAMSESHXIKTLQGDNDVPGPDSRLPAVQENQAEPRVILDSFALIKI 107
DB 23 LMDQPELILNESN-----DQANWEN--ITLIL-CASRSIRSKTLLKLP 113
OY 75 -----PAIKH---OFI---LTGDPGQRYGRSGSTGQWQMSLILATPC 113
DB 66 TQMTWIPRSHRTQVSPLIGALTSTHAGLVRCQWKEBTQMSKPSVLELEAP 117

```


RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK001427; BA01685.1; -
 SO SEQUENCE 458 AA: 50748 MW: 493594D2C1B4CD7 CRC64:

Query Match 11.3% Score 75; DB 4; Length 458;
 Best Local Similarity 31.5% Pred. No. 3.6;
 Matches 29; Conservative 10; Mismatches 23; Indels 30; Gaps 6;

OY 33 ASSEHXLKTLQCCDADVPGRPG-----DSRLPAVGEAGQGVYHIDSP-AIKHQ 80
 DB 3 ADEEVYLNILAECDMLGRPGSTFSQKGRABSR-----GSQSRHLLSPOLVQYO 55
 DB 81 FLITGDT--GGRYRCRSGL-----STGMQOL 104
 DB 56 ---CGDSGKQGRRTSRSSCKCCCLEMNMWSQL 84

RESULT 14
 ID 09BVM3 PRELIMINARY: PRT: 458 AA.

AC 09BVM3: 01-JUN-2001 (TREMBALEL. 17, Created)
 DT 01-JUN-2001 (TREMBALEL. 17, Last sequence update)
 DE 01-JUN-2001 (TREMBALEL. 17, Last annotation update)
 DE HYPOTHEICAL 50.7 KDA PROTEIN.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA STRAUSBERG R.:
 RC TISSUE-LUNG CARCINOMA;
 RA Strausberg R.: 2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AK001427; BA01685.1; -
 SO SEQUENCE 458 AA: 50706 MW: 492984CEADCF50A7 CRC64:

Query Match 11.3% Score 75; DB 4; Length 458;
 Best Local Similarity 31.5% Pred. No. 3.6;
 Matches 29; Conservative 10; Mismatches 23; Indels 30; Gaps 6;

OY 33 ASSEHXLKTLQCCDADVPGRPG-----DSRLPAVGEAGQGVYHIDSP-AIKHQ 80
 DB 3 ADEEVYLNILAECDMLGRPGSTFSQKGRABSR-----GSQSRHLLSPOLVQYO 55
 OY 81 FLITGDT--GGRYRCRSGL-----STGMQOL 104
 DB 56 ---CGDSGKQGRRTSRSSCKCCCLEMNMWSQL 84

RESULT 15
 ID 09BVM3 PRELIMINARY: PRT: 647 AA.

AC 09BVM3: 01-JUN-2001 (TREMBALEL. 17, Created)
 DT 01-JUN-2001 (TREMBALEL. 17, Last sequence update)
 DE 01-JUN-2001 (TREMBALEL. 17, Last annotation update)
 DE HYPOTHEICAL 71.7 KDA PROTEIN.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA STRAUSBERG R.:
 RC TISSUE-CHORIOCARCINOMA;
 RA Strausberg R.: 2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC000132; AAH00132.1; -
 SO SEQUENCE 647 AA: 71717 MW: 945388811311460 CRC64:

Query Match 11.3% Score 75; DB 4; Length 647;
 Best Local Similarity 31.5% Pred. No. 5.2;
 Matches 29; Conservative 10; Mismatches 23; Indels 30; Gaps 6;

OY 33 ASSEHXLKTLQCCDADVPGRPG-----DSRLPAVGEAGQGVYHIDSP-AIKHQ 80
 DB 3 ADEEVYLNILAECDMLGRPGSTFSQKGRABSR-----GSQSRHLLSPOLVQYO 245
 DB 81 FLITGDT--GGRYRCRSGL-----STGMQOL 104
 DB 245 ---CGDSGKQGRRTSRSSCKCCCLEMNMWSQL 273

RESULT 16
 ID 09P209 PRELIMINARY: PRT: 648 AA.

AC 09P209: 01-OCT-2000 (TREMBALEL. 15, Created)
 DT 01-OCT-2000 (TREMBALEL. 17, Last sequence update)
 DE 01-JUN-2001 (TREMBALEL. 17, Last annotation update)
 DE KIA1519 PROTEIN (FRAGMENT).

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RA MEDLINE:20277482; Pubmed 10819331;
 RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.:
 RT Prediction of the coding sequences of unidentified human
 RT genes XVII. The complete sequences of 100 new cDNA clones from brain
 RT match code for large proteins in vitro. *;
 DR EMBL: AF040952.1; AF040953.1; -
 DR InterPro: IPR001611; LRR.

DR Pfam: PF00560; LRR: 2.
 DR SMART: SM00446; LRRcap: 1.
 FT NON_TER
 SO SEQUENCE 648 AA: 71801 MW: 44AC3750072B661E CRC64:

Query Match 11.3% Score 75; DB 4; Length 648;
 Best Local Similarity 31.5% Pred. No. 5.2;
 Matches 29; Conservative 10; Mismatches 23; Indels 30; Gaps 6;

OY 33 ASSEHXLKTLQCCDADVPGRPG-----DSRLPAVGEAGQGVYHIDSP-AIKHQ 80
 DB 193 ADEEVYLNILAECDMLGRPGSTFSQKGRABSR-----GSQSRHLLSPOLVQYO 245
 OY 81 FLITGDT--GGRYRCRSGL-----STGMQOL 104
 DB 246 ---CGDSGKQGRRTSRSSCKCCCLEMNMWSQL 274

RESULT 17
 ID 070434 PRELIMINARY: PRT: 663 AA.

AC 070434: 01-AUG-1998 (TREMBALEL. 07, Created)
 DT 01-AUG-1998 (TREMBALEL. 07, Last sequence update)
 DE 01-JUN-2001 (TREMBALEL. 17, Last annotation update)
 DE KILLER ACTIVATOR RECEPTOR-LIKE PROTEIN P91D (FRAGMENT).

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN-B10.A: TISSUE-PERITONEUM;
RA MEDLINE-98218758; PubMed-9538215;
RA Yamashita Y., Fukuta D., Tsuji A., Nagabukuro A., Matsuda Y.,
RA Nishikawa Y., Ohyama Y., Ohmori H., Ono M., Takai T.;
RT Genomic structures and chromosomal location of p51, a novel murine
RT regulatory factor, p51-358 (1998).
CC J Biol Chem 273:358-368(1998).
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AF055896; AAC40073.1; -.
DR HSSP: P43626; 1MKR.
DR MGD: MGI:1276541; P1ra11.
DR InterPro: IPR003209; 19_1.Like.
DR InterPro: IPR003006; 19_1.Like.
DR Pfam: PF00047; 19_1; 5.
DR SMART: SM00409; 1G_1; 3.
DR SMART: SM00410; 1G_1Like; 2.
FT NON_TER 1
FT SEQUENCE 663 AA; 73761 MW; 659886E10BABB95 CRC64;
SO QUERY MATCH 11.3%; Score 74.5; DB 11; Length 663;
Best Local Similarity 28.6%; Pred. No. 6.1;
Matches 32; Conservative 11; Mismatches 48; Indels 21; Gaps 4;
OY 6 VELLINGWGPVTEA---AIFETQXSLAESBHLKLTGCCADADVPQPGDSRLP 61
DB 215 VELLVSGNLOKPTKASPGSYTSKANTLMQO-----GNLDAY-----YFLINE 260
OY 62 QEMKQEPVHDSPAIKHPIILDTGQ---GRYKRSGLSTQXLSKLTLE 110
DB 261 KSKQTSQTOYLOQPGNKGKFPISVYQDHAGYRCGYASGWSQSPDITL 312
RESULT 18
OY055001 PRELIMINARY; PRT: 680 AA.
AC 055001:
DT 01-JUN-1998 (TREMURel. 06, Created)
DT 01-JUN-1998 (TREMURel. 06, Last sequence update)
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)
DE PAIRED-IG-LIKE RECEPTOR A10 (KILLER CELL INHIBITOR RECEPTOR-LIKE
DE RECEPTOR FAMILY).
CC PIRATIN (P1B).
CC Mus musculus (Mouse).
CC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC NCBI_Taxid-10090;
CC (1)
CC SEQUENCE FROM N.A.
CC STRAIN-23/5V; Fukuta D., Tsuji A., Nagabukuro A., Matsuda Y.,
CC Nishikawa Y., Ohyama Y., Ohmori H., Ono M., Takai T.;
CC J. Biochem. 123:0-0(1998).
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AF041035; AB96927.1; -.
DR HSSP: P43626; 1MKR.
DR MGD: MGI:1276541; P1ra10.
DR InterPro: IPR003599; 19_1.Like.
DR InterPro: IPR003600; 19_1.Like.
DR InterPro: IPR003006; 19_1.Like.
DR Pfam: PF00047; 19_1; 5.
DR SMART: SM00409; 1G_1; 3.
DR SMART: SM00410; 1G_1Like; 2.
DR SEQUENCE 680 AA; 75597 MW; 531DPE6D291BE11 CRC64;
SO QUERY MATCH 11.3%; Score 74.5; DB 11; Length 680;
Best Local Similarity 28.6%; Pred. No. 6.2;
Matches 32; Conservative 11; Mismatches 48; Indels 21; Gaps 4;

```

```

OY 6 VELLINGWGPVTEA---AIFETQXSLAESBHLKLTGCCADADVPQPGDSRLP 61
DB 215 VELLVSGNLOKPTKASPGSYTSKANTLMQO-----GNLDAY-----YFLINE 260
OY 62 QEMKQEPVHDSPAIKHPIILDTGQ---GRYKRSGLSTQXLSKLTLE 110
DB 261 KSKQTSQTOYLOQPGNKGKFPISVYQDHAGYRCGYASGWSQSPDITL 312
RESULT 19
OY05018 PRELIMINARY; PRT: 192 AA.
AC 05018:
DT 01-NOV-1998 (TREMURel. 12, Created)
DT 01-NOV-1998 (TREMURel. 12, Last sequence update)
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)
DE VIF PROTEIN.
CC VIF.
CC Human immunodeficiency virus type 1.
CC Viruses; Retroid viruses; Retroviridae; Lentivirinae.
CC NCBI_Taxid-11676;
CC SEQUENCE FROM N.A.
CC STRAIN-V108003ALT;
CC Hassane L., Agostini I., Candotti D., Bessou G., Caballero M.,
CC Ayut H., Autran B., Barthalay Y., The French ALT Study Group.
CC Vigne R.;
CC Characterisation of human immunodeficiency virus type 1 vif gene in
CC long term asymptomatic individuals.
CC RT (1998).
CC EMBL: AF143108; AB037875.1; -/db/embank/dbj databases.
CC InterPro: IPR00475; Viral_Infect.
CC Pfam: PF00559; VIF_1.
CC PRINTS: PR00349; VIRIONINFECT.
CC PRODOM: PD00063; Viral_Infect. 1.
CC SEQUENCE 192 AA; 22805 MW; CBB7D94B976F6280 CRC64;
SO QUERY MATCH 11.2%; Score 74; DB 12; Length 192;
Best Local Similarity 22.0%; Pred. No. 1.8;
Matches 22; Conservative 15; Mismatches 39; Indels 24; Gaps 3;
OY 8 LLUCVY-----TWSPTEAIPFETQXSLAESBHLKLTGCCADADVPQPGDSRLP 61
DB 8 WLVQVDRKRTKMSLKHITSLKSKAKQYKRRHETSPASVHPIPLDANLVIT 67
OY 62 QEMKQEPVHDSPAIKHPIILDTGQ---GRYKRSGLSTQXLSKLTLE 110
DB 68 TWVG-----LITGERE--WHLGQGVSEFM 89
RESULT 20
OY05018 PRELIMINARY; PRT: 289 AA.
AC 075018:
DT 01-NOV-1998 (TREMURel. 08, Created)
DT 01-NOV-1998 (TREMURel. 08, Last sequence update)
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)
DE LENCOCYTE IMMUNOGLOBULIN-LIKE RECEPTOR-6B.
CC LTR-6.
CC Haplorhina; Hominidae; Homosapiens (human).
CC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_Taxid-9606;
CC (1)
CC SEQUENCE FROM N.A.
CC Borges L., Hsu M.-H., Ronger N., Rubin M., Cosman D.;
CC J. Immunol. 160:0-0(1997).
CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AF025529; AB87663.1; -.
DR HSSP: P43626; 1MKR.
DR InterPro: IPR003600; 19_1.Like.

```


DR PRINTS: PRO0049; VIRIONINFECT.
 DR PRODM: PRO00063; VIRAL_INFECTION.
 SQ SEQUENCE 192 AA: 22651 MW: 02612B2B5506598C CRC64:

Query Match 10.6% Score 70; DB 12; Length 192;

Best Local Similarity 19.7% Pred. No. 5;

Matches 24; Conservative 19; Mismatches 43; Indels 36; Gaps 4;

OY 8 LLLKCV-----TWGPTVALFYETQXSXLMASEHXKLTLGGCDADVPQPGDSRLAP 61

DB 8 LTVWQDPRMKIRTNMISLVKHNHYSRASCQFYRHNRESRHPRVSAKVNHLGDAVLTK 67

OY 63 QEWGAQEPVHDSPAIKHOFILTDQGRACRQSCSLSTQKXLSKLTETGPKVYACSLA 121

DB 68 TWYGIQ-----TGERE--MHLNGVSIEM-----RLRCSQ 97

OY 122 LD 123

DB 98 VD 99

RESULT 40

OY94SO PRELIMINARY: PRT: 192 AA.

AC 09494SO: 01-JUN-2001 (TREMURel. 17, Created)

DT 01-JUN-2001 (TREMURel. 17, Last sequence update)

DE 01-JUN-2001 (TREMURel. 17, Last annotation update)

DE VIF PROTEIN.

GN Human Immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OK NCBI_TaxID=11676;

RP SEQUENCE FROM N.A.

RC STRAIN=94IN176;

RA MEDLINE=21094715; PubMed=1177395;

RA Rodenburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,

RA Kallish M.L., Shaw G.M., Allen S., Hahn B.H., Gao P.;

RT Near full-length clones and reference sequences for subtype C

RT isolates for HIV type 1 from three different continents.;

RL AIDS Res. Hum. Retroviruses 17:101-108(2001).

RP SEQUENCE FROM N.A.

RC STRAIN=94IN176;

RA Rodenburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,

RA Allen S., Shaw G.M., Hahn B.H., Gao P.;

RT Submitted (01-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL=AF040952; ANK535211;

SQ SEQUENCE 192 AA: 22625 MW: EDFGCI5F8211645 CRC64:

Query Match 10.6% Score 70; DB 12; Length 192;

Best Local Similarity 27.3% Pred. No. 5;

Matches 18; Conservative 10; Mismatches 32; Indels 6; Gaps 1;

OY 8 LLLKCV-----TWGPTVALFYETQXSXLMASEHXKLTLGGCDADVPQPGDSRLAP 61

DB 8 LTVWQDPRMKIRTNMISLVKHNHYSRASCQFYRHNRESRHPRVSAKVNHLGDAVLTK 67

OY 62 QEWGAQ 67

DB 68 TWYGIQ 73

RESULT 41

OY94SO PRELIMINARY: PRT: 728 AA.

AC 09494SO: 01-MAY-1999 (TREMURel. 10, Created)

DT 01-MAY-1999 (TREMURel. 10, Last sequence update)

DE 01-MAY-1999 (TREMURel. 10, Last annotation update)

DE COS54-22.
 GN Herpesvirus 1 (Lucke tumor herpesvirus).
 OS Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Viruses; Retroid viruses; no RNA stage; Herpesviridae.
 OK NCBI_TaxID=85655;

RP SEQUENCE FROM N.A.

RC STRAIN=94IN176;

RA Davisson A.J., Sauerbier W., Dolan A., Addison C., McKinnell R.G.;

RT Genomic studies of the Lucke tumor herpesvirus (RaV1-1).;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RL EMBL=AF111004; ABD1238511;

SQ SEQUENCE 728 AA: 81566 MW: 76P7921CA573743D CRC64:

Query Match 10.5% Score 69.5; DB 12; Length 728;

Best Local Similarity 25.2% Pred. No. 24;

Matches 30; Conservative 11; Mismatches 59; Indels 19; Gaps 3;

OY 18 VYBAIRIYETQXSXLMASEHXKLTLGGCDADVPQPGDS-----RLPQVQENQAE 68

DB 250 YCAAGAGYIIIPWLFRATYKHXKLTGCDALVQYGGEGMLDEKQKTLRLPAREENGTSP 209

OY 63 PVHDSPAIKHOFILTDQGRACRQSCSLSTQKXLSKLTETGPKVYACSLAD 123

DB 310 PALLTCDDVPASVYICDHPDLSRMSLTFQAGACQCN-----EVCVGRVYNTLLD 362

RESULT 42

OY94SO PRELIMINARY: PRT: 841 AA.

AC 054999: 01-JUN-1998 (TREMURel. 06, Created)

DT 01-JUN-1998 (TREMURel. 06, Last sequence update)

DE 01-JUN-2001 (TREMURel. 17, Last annotation update)

DE PROTEIN P21A).

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RC STRAIN=125/SV.

RA Takai T., Yamashita Y., Fukuta D., Tsuji A., Nagabukuro A.;

RA Matsuda Y., Nishikawa Y., Ohyama Y., Ohmori H., Ono M.;

RL J. Biochem. 123:0-0(1998).

RP SEQUENCE FROM N.A.

RC STRAIN=125/SV.

RA Takai T., Yamashita Y., Fukuta D., Tsuji A., Nagabukuro A.;

RL J. Biochem. 123:0-0(1998).

RP SEQUENCE FROM N.A.

RC STRAIN=125/SV.

Query Match 10.5% Score 69.5; DB 11; Length 841;

Best Local Similarity 27.2% Pred. No. 28;

Matches 31; Conservative 13; Mismatches 45; Indels 25; Gaps 5;

OY 6 VYBAIRIYETQXSXLMASEHXKLTLGGCDADVPQPGDSRLAP 61

DB 8 LTVWQDPRMKIRTNMISLVKHNHYSRASCQFYRHNRESRHPRVSAKVNHLGDAVLTK 67

OY 62 QEWGAQ 67

DB 68 TWYGIQ 73

RESULT 41

OY94SO PRELIMINARY: PRT: 728 AA.

AC 09494SO: 01-MAY-1999 (TREMURel. 10, Created)

DT 01-MAY-1999 (TREMURel. 10, Last sequence update)

DE 01-MAY-1999 (TREMURel. 10, Last annotation update)

```

Db      215 VELLASGNLQKPFIRKAEVSIVTSKRAATIMQ-----GMLAYFALHNGSGOKTO 266
Oy      60 AVEGKGOEPVHLDSPAIKHOFLITLTDGQ---GRYKRSGLSTGXOLSKLEL 110
Db      267 STGT-----LQDGNKGRKFTFSPHTRHGMQVRCYCSGAKMSQSPDLLE 312

RESULT 43
071162          PRELIMINARY:      PRT: 192 AA.
AC 071162: 1998 (TREMblrel. 07, Created)
AC 071162: 1998 (TREMblrel. 07, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VIF PROTEIN.
GN VIF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
RC STRAIN=C18;
RA Oelrichs R.B., McPhee D.A., Deacon N.J.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF042102; AAD03210.1;
DR InterPro: IPR000475; Viral_infect.
DR Pfam: PF00559; Vif; 1 Viral_infect.
DR PRINTS: PR00063; VIRIONINFECT.
DR PRODOM: PD00063; Viral_infect. 1.
SQ SEQUENCE 192 AA: 22500 MW: 22A07D251EC1ED40 CRC64:

```

```

Query Match      10.4%: Score 69; DB 12; Length 192;
Best Local Similarity 20.6%: Pred. No. 6.4;
Matches 21; Conservative 16; Mismatches 41; Indels 24; Gaps 3;

Oy      8 LLLKGV-----TWGPTVEAIFYETQXSLMAESKHILTKLGGCDADVGPFGDSRLPAY 61
Db      8 MIVQGVDRMRIRTWMSLVKHHYVSKKACMYRHHENTHPRKISSVHPIGLDARLVYT 67
Oy      62 QEMGKGOEPVHLDSPAIKHOFLITLTDGQRCRSGSLSTGXMO 103
Db      68 TWG-----LITCERE--MHGNGVSIEMKRS 91

RESULT 44
071166          PRELIMINARY:      PRT: 192 AA.
AC 071166: 1998 (TREMblrel. 07, Created)
AC 071166: 1998 (TREMblrel. 07, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VIF PROTEIN.
GN VIF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
RC STRAIN=FROM N.A.
RA Mwaengo D.M., Novembre F.J.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF049494; AAC68843.1;
DR InterPro: IPR000475; Viral_infect.
DR Pfam: PF00559; Vif; 1.
DR PRINTS: PR00349; VIRIONINFECT.
DR PRODOM: PD00063; Viral_infect. 1.
SQ SEQUENCE 192 AA: 22559 MW: 00DE1456317A0673 CRC64:

```

Query Match 10.4%; Score 69; DB 12; Length 192;
Best Local Similarity 20.2%; Pred. No. 6.4;

```

Matches 21; Conservative 18; Mismatches 41; Indels 24; Gaps 3;

Oy      8 LLLKGV-----TWGPTVEAIFYETQXSLMAESKHILTKLGGCDADVGPFGDSRLPAY 61
Db      8 MIVQGVDRMRIRTWMSLVKHHYVSKKACMYRHHENTHPRKISSVHPIGLDARLVYT 67
Oy      62 QEMGKGOEPVHLDSPAIKHOFLITLTDGQRCRSGSLSTGXMO 105
Db      68 TWG-----LITCERE--MHGNGVSIEMKRS 93

RESULT 45
071171          PRELIMINARY:      PRT: 192 AA.
AC 071171: 1997 (TREMblrel. 07, Created)
AC 071171: 1997 (TREMblrel. 07, Last sequence update)
DE 01-AUG-1998 (TREMblrel. 07, Last annotation update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VIF PROTEIN.
GN VIF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
RC STRAIN=FROM N.A.
RA Mwaengo D.M., Novembre F.J.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF04138200.475; AAC68852.1;
DR InterPro: IPR000475; Viral_infect.
DR Pfam: PF00559; Vif; 1 Viral_infect.
DR PRINTS: PR00349; VIRIONINFECT.
DR PRODOM: PD00063; Viral_infect. 1.
SQ SEQUENCE 192 AA: 22398 MW: A6E790B042ABC996 CRC64:

```

```

Query Match      10.4%: Score 69; DB 12; Length 192;
Best Local Similarity 20.2%: Pred. No. 6.4;
Matches 21; Conservative 18; Mismatches 41; Indels 24; Gaps 3;

Oy      8 LLLKGV-----TWGPTVEAIFYETQXSLMAESKHILTKLGGCDADVGPFGDSRLPAY 61
Db      8 MIVQGVDRMRIRTWMSLVKHHYVSKKACMYRHHENTHPRKISSVHPIGLDARLVYT 67
Oy      62 QEMGKGOEPVHLDSPAIKHOFLITLTDGQRCRSGSLSTGXMO 105
Db      68 TWG-----LITCERE--MHGNGVSIEMKRS 93

Search completed: January 7, 2002, 16:51:26
Job time: 273 sec

```

